AD-A007 963

SIMULATION AND STUDY OF HIGH WORKLOAD OPERATIONS

Advisory Group for Aerospace Research and Development

April 1974

AD-A007 963

AGARD-CP-146

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.146

SIMULATION AND STUDY OF HIGH WORKLOAD OPERATIONS

Edited by

Wing Commander A.N.Nicholson, RAF

RAF Institute of Aviation Medicine
Farnborough, Hants
United Kingdom

PRICES SUBJECT TO CHANGE

Papers presented at the Aerospace Medical Panel Specialists Meeting,
held at Oslo, Norway, 24–25 April 1974.

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Program and the Aerospace Applications Studies Program. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced
directly from material supplied by AGARD or the author.

Published October 1974

355.469:613.6:613.73

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London, W1P 1HD

AEROSPACE MEDICAL PANEL

Panel Chairman:	Colonel J.F.Culver, USAF, MC
Panel Deputy Chairman:	Major General H.S.Fuchs, GAF, MC
Panel Executive:	Lt Colonel P.Varène, FAF

MEETING ORGANIZATION

Host Coordinator:	Lt Colonel A.Borg, RNoAF
Technical Program Organizer:	Wing Commander A.N.Nicholson, RAF

PREFACE

The use of simulation for the evaluation of control dynamics, system components and procedures is well established, less certain is the use of such techniques in the evaluation of aircrew performance in high workload situations of an operational nature. Studies in the field present considerable difficulties, particularly when factors which may adversely affect flight safety have to be assessed. Although laboratory and airborne simulation of an operational situation may lack reality, such techniques may be the only way to predict operational performance.

The meeting was intended to provide a series of presentations including analyses of operational aspects from several NATO countries. It was hoped that recent studies on high workload situations in aerospace operations and accounts of simulation of such task conditions in the laboratory and in flight would be included.

CONTENTS

* Paper not available for publication.

TECHNICAL EVALUATION

The meeting evaluated the contribution which simulation and study of operations could make to the understanding of intensive air operations. It was evident from the discussions that, though simulation provided opportunities to examine in detail limited aspects of air operations, it nevertheless lacked the reality essential for a definitive analysis. On the other hand the nature of actual operations often precluded a full analysis of the problem.

Several aspects of methodology were discussed. An important technique in the analysis of high workload was that of subjective assessment, but it is a technique which needs careful study. Controls for subjective assessments were frequently obtained under conditions of reasonable workload, and the question arose whether subjects used the same technique to evaluate workload under the circumstances of an intensive operation. There was a need for subjects to assess the individual factors which contributed to overall workload, and in this way it may be possible to create models of workload assessment and to define the stress which leads to a breakdown of the model. With subjective assessments of workload the use of other parameters, such as physiological change, should be considered. Physiological changes and assessments of workload should correlate with each other, but if they are dissociated a careful assessment of the problem is needed.

Another area of importance was the use of physiological parameters as measures of workload. There was no real evidence that physiological change was a measure of workload *per se*, but it may indicate a change in the physiological balance of the body brought about by intensive workload. Change in physiological balance is a response to workload, but it may also tell us how the individual adapts and suggest when adaptation is breaking down. The breakdown of the ability to adapt may be associated with impaired performance. Several studies were presented on the response of the endocrine system to stress and it would appear that advances are being made in this direction. There is a need for a much more basic understanding of endocrine physiology under stress and its application to intensive military operations.

The relation of optimum workload and optimum performance was discussed. Pilot ability during overload may not be optimum, and it is in this direction that simulation studies may be able to play a greater part. The problem of simulating high workload is a difficult one, and the use of simulation for studying high workload operations is an area of uncertainty. But simulation may have a part to play as it is an experimental situation where systems engineers, physiologists, psychologists and aircrew can work together.

A further area of interest was the prediction of operational capability. In this area aircrew to aircraft ratios, continuous operational capability and preservation of skill under stress were included. If we were able to predict operational capability then a considerable step would have been made in the human factors aspect of intensive military operations. It is in this area where future work needs to be carried out.

A.N. Nicholson.

A.N.NICHOLSON
Wing Commander
RAF Institute of Aviation Medicine
Farnborough, Hampshire, UK

Deputy Chairman, Behavioural Sciences
Committee, ASMP

SYSTEMS SIMULATION: A GLOBAL APPROACH TO AIRCREW WORKLOAD

by

Harry M. Hughes, Ph.D.
Bryce O. Hartman, Ph.D.
Raul Garcia, M.S. (BRP)
Paul Lozano, B.A. (BRP)
USAF School of Aerospace Medicine
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas, USA 78235

ABSTRACT

Aircrew workload can be studied at many different levels of detail. In the most general sense, it is a function of the total workload imposed upon a unit in relation to the number of crews in that unit. An airlift system simulation program has been designed using this global approach and a number of simulation studies have been performed. Outcomes in terms of systems effectiveness measures, crew workload, and crew welfare measures will be presented.

INTRODUCTION

A general simulation computer program capable of representing the major operational attributes of a typical squadron of MAC jet transport aircraft and aircrewmen has been developed by the School of Aerospace Medicine. Given the resources (how many planes, crews), the workload (number and route of missions), and the rules under which to operate (various regulations), the program schedules the missions, selects the crews and planes, and flies the missions, inserting random fluctuations to represent delays and weather variations. During the course of the simulation, the program tracks how the system is performing by acquiring operational data such as cancellations, flying time on each leg, and delays, which can be later assembled into such statistics as time away from home and work-month hours. The following sections will describe in some detail the general mechanics of the simulation program indicating the input requirements (what we decide beforehand), the simulation logic (what happens while it is running), and the output measures generated by each simulation run (what can be analyzed afterward). A typical use of the program will be described and some results will be reported, showing the application of the program in making management decisions affecting aircrews.

DESCRIPTION OF THE SIMULATION PROGRAM

Input Requirements (what we decide beforehand)

The information to be supplied to the simulation prior to a run can be grouped into five general categories: policy, route system characteristics, resources, workload, and predestined events. The first three of these are placed in the computer at the beginning of the run; the last two are prepared prior to the run and placed on a tape to be fed in as their dates and times come up during the run. We will describe these two groups of information under the headings "Initialization" and "Exogenous Events" respectively.

a. Initialization.

The first group of input parameters provides the means by which policy rules are established. Some of the more important policy input parameters are:

1. Crew Type--The crew types which we have thus far provided are either "basic" (with five members flying according to regulations), "augmented" (with a complement of up to ten men on duty at the same time), or "double" (with five men on duty at any given time while a second shift of five men is aboard in a resting status). This latter type, which alternates crews from duty to rest during the entire mission, is one of the suggested policies that this simulation was developed to investigate. The initialization defines the number of pilots, navigators, flight engineers and loadmasters in the squadron. These crew types can be simulated without further computer programming; other crew types could be introduced with rather simple additional programming.

2. Factor for Awarding Free Time--At the completion of a mission, free time is awarded by multiplying the time away from home by a factor set during initialization. A maximum limit for the amount of free time awarded at any one time is also established at initialization. Free time based on differing credits for differing kinds of duty time would require additional programming.

3. Maximum Time Allowed for Delay at Home Base--This parameter is used in determining when to reschedule or cancel a mission. At bases other than home base, missions are not re-scheduled or cancelled, but may be further delayed if the maximum duty day initially prescribed for that mission is exceeded.

4. Maximum Flying Time Limits--The simulation program provides for two kinds of limitations on maximum flying time per individual per period. Definition of such periods is quite flexible.

5. Length of Crew Rest--This is the number of hours which policy states that a crew shall rest prior to the mission departure. The same or another figure is specified for crew rest enroute upon arrival at a stage base or upon expiration of a duty day.

6. Air Time Between Aircraft Maintenance--While the principal emphasis of this simulation lies on crew data and crew affects, provision has been made to take aircraft out of service at home base for two different lengths of periods called minor maintenance and major maintenance. This input parameter specifies the number of hours of air time that shall not be exceeded before minor maintenance, the number of hours the plane will be out of service for minor maintenance, the number of air hours that shall not be exceeded between periods of major maintenance, and the number of days required for major maintenance. The simulation program during the run will utilize the expected air time of any proposed mission to estimate whether a particular aircraft would exceed its maintenance limitations and will accept or reject the aircraft for that mission on that basis. Thus, the interactive affects of periodic maintenance limitations upon crews are taken into account. A simple modification also provides the ability to simulate isochronal maintenance, taking planes out of service at fixed periods of calendar time irrespective of their air time.

7. Number of Stage Crews at Each Base--Part of the initialization is to place the number of stage crews needed at each of the stage points according to policy to be used in that run. This policy is usually expressed as a ratio such as 1/30, meaning one stage crew for each 30 missions per month planned to move through that stage pool.

Having looked at some of the major initialization parameters which specify policy we will now describe some which specify the system characteristics.

8. Route Types--This set of parameters specifies for each route type the number of legs, the direction of each leg (in-bound or out-bound), scheduled or average air times of each leg, and identification of the bases. The leg direction is pertinent to the setting up of directional stage crew pools. Note that the inclusion of scheduled air times essentially identifies one type of aircraft with a route type. For a different type of aircraft over the same geographic route, an additional route type is defined to the simulation.

9. Mission Type--For each kind of mission to be flown a code is entered identifying that type of mission, the crew type to be used on that mission, the route type of the mission, the bases where staging will occur, the total estimated air time and the maximum duty day for each leg. Note that this does not specify any dates of departure or arrival but merely describes the characteristics of a particular type of mission. Later during the course of the run, the exogenous tape will indicate several times during the course of a month that a mission of type so-and-so should be launched at such-and-such a time on such-and-such a day. That information gives the identifying code and time of launch; the simulation will then refer to the initialized mission type information to find out all the details involved in the setup.

10. Pre-Flight Length and Delays--In actual practice and, hence, in the simulation, the length of time in pre-flight has a normal planned value which is initially furnished here, but for various causes occurring in fairly random fashion actual take-offs are delayed. Excluding the nonavailability of plane or crew, we have made provision for all other delays to be included in a random distribution which is initially submitted at this point. Later during run time as each leg is about to be launched, the simulation gets a random sample from this specified distribution and sets the actual departure time accordingly.

11. Weather and Other Variabilities in the Air--In order to provide for the random fluctuations in length of air time along a particular leg in a particular direction brought about operationally by fluctuation in such things as wind, power settings, weights and a miscellany of other factors, we provide for an initial specification of a distribution of factors to be applied to the average no-wind-time for any leg.

Finally, in the initialization phase we specify the resources that a particular run will have available. They include the following parameters:

12. Number of Flight-Qualified Personnel--The number of pilots, co-pilots, navigators, engineers, and loadmasters in the squadron is specified separately and need not be equal. Within each of these groups all personnel are presumed to be flight-qualified and no provision has thus far been made for trainees and examiners. It is quite feasible with a small amount of additional programming to include other types of personnel. The crew type parameter described above would then have to be supplemented so that at least one crew type calls for one of the new type of personnel.

13. Number of Planes--The total number of planes assigned to the squadron is initially entered.

b. Exogenous Events. The second type of information which is prepared in advance of a run is an exogenous event tape. This tape is a calendar of events due to occur at a time independent of what goes on in the simulation. It includes pre-planned calls upon the squadron for missions, and it provides a mechanism for handling sick leave, planned leave, emergency leave, general military duties, and special time-clocked squadron duties.

1. Mission Workload--The mission workload is placed on the exogenous tape in the form of a notice to the squadron, each mission notice received just in time to designate a crew and aircraft for that mission.

2. Unscheduled Leave--This is provided by notices at random times that a particular person

becomes sick or has an emergency call for leave. In order to generate these notices we sample a random distribution which on the average will give us the experience incurred operationally as to frequency of emergency calls, of falling ill and the length of the ill period. The results of this sampling are then placed on the exogenous tape in order of occurrence so that during the run the system will find out about each occurrence only when it happens.

3. Adjustable Duties--Those military duties included in manpower analyses under the heading "AFNAT," which stands for Air Force Non-available Time and includes ordinary leave, are spread out over the month and assigned to each individual at specific times on the exogenous tape. The difference from the foregoing categories is that during the run a crewman, called upon for this type of activity while away from home base, will have it saved for him until he returns and completes his free time; then he will be assigned to these duties.

4. Fixed Duties--The third type of predetermined assignment for an individual is intended to cover such items as squadron alpha alert, ground training, and scheduling officer duties. These events occur not randomly but regularly and must be fulfilled as scheduled, so that during run time our computer looks ahead and will not select a crewman for a particular mission if the forecast length of that mission would conflict with his being back in time for a duty specified here.

Thus the exogenous tape provides the simulation program with a timing sequence for setting up missions and for putting crew members on leave or special duties.

Simulation Logic (what happens while it is running)

A simulation run is begun by loading the program, mounting the exogenous tape, and reading in the initialization values. This includes assembling the required initial stage crews and pre-staging them out in the system. All planes and remaining crewmen are placed in pools of available planes and crewmen. The simulation clock is started by reading in the first notice from the exogenous tape. As mentioned earlier, the exogenous tape contains notices which schedule a mission or notices to place an individual man on some "blocked out" status (making him unavailable for a mission for a specified period). We will illustrate the detailed events involved, by tracing a single mission from start to finish. We will discuss first those events which occur at home base prior to mission departure, then the events which take place at an enroute base, and finally those events which occur at the completion of the mission. The computer actually performs the actions for all missions by the simulator clock so that it may successively select a crew for tomorrow's mission launch, alert a stage crew at base F for an imminent departure, take a crew [illegible] of free time status, launch an in-bound mission from an intermediate base, launch today's noon missi[illegible] from home base, "arrive" yesterday's mission at an enroute base and compute its subsequent time of launch from the delay distribution, all in order of the clocktime at which each event is supposed to occur.

a. Pre-departure Events

Notices to schedule a mission have been placed on the exogenous tape so that the notice appears on the calendar (the computer finds out about it) 24 hours before the scheduled departure time for the mission. At this time, selection of the plane and crew begins. Information about crew-type, route, number of staging bases, and the length of time the crew will be away from home is obtained from the input. This information is used to insure that policy rules are not violated in selecting the crew and plane. A crewman's eligibility is determined by:

1. Is he available? A crewman is not assigned to a mission which will conflict with his fixed duties. If a conflict inadvertently occurs because the mission actually lasts longer than the normal time, the duties will be performed at the completion of the mission. He is not available if he is on unscheduled leave (sick, emergency). He may be assigned to a mission even if it conflicts with adjustable duties, which are then carried out after completion of the mission.

2. At the beginning of each leg, maximum flying time per crewman is imposed, for both short (normally 30-day) and long (normally 90-day) periods, without waiver except to complete a leg whose normal length would not have put him over the limit. In making up crews at home base, an attempt is made to see that each crewman selected can complete the entire mission without exceeding his short or long flying time maximums. If no one of a particular position such as flight engineer meets this criterion, then only the flying time to the first staging point is used to determine availability.

3. Who goes first?--If more than one crewman of a given position is available, the one with the least accumulated flying time for the quarter (long period) is selected. In case of ties, the one with the least accumulated flying time for the month (short period) is selected.

A plane's eligibility is determined by:

1. Will the plane be at home base in time to begin pre-flight?

2. Will the sum of the plane's accumulated air time and the air time required for this mission exceed the time for the next minor or major maintenance for that aircraft? Or alternatively, will the proposed mission departure time plus normal mission elapsed time interfere with a scheduled isochronal maintenance for that plane?

If there is either no plane or no crew for a particular mission, the mission is cancelled. However, the mission is merely rescheduled to depart at the earliest possible time if a plane and crew can be found so that such rescheduling will not exceed the maximum time (supplied in the initialization) allowed for a delay at home base. As soon as selection of the crew is completed, they are placed on home crew rest. Upon completing crew rest (usually 12 hours), they are allowed travel time to report to the base. The length of this status is currently fixed at one hour. Normally, preflight of the plane begins at this time. It is possible, however, for the plane assigned to this mission to be unavailable. This can occur if the plane assigned is still in maintenance, but will be available in time to prevent a major delay. If the crew has to wait, they are placed in a status called ramp time. As soon as both plane and crew are available, pre-flight status begins for all. As stated earlier, the actual ground time is a random value. If the actual ground time does not exceed the scheduled ground time, the plane departs and the crew status changes to flying time which is charged against their monthly and quarterly limits. If the scheduled ground time is exceeded, then the crew is placed in ramp status (corresponds to ramp pounding due to unscheduled maintenance or weather delays). The crew is allowed to depart if the ramp time does not exceed six hours (a figure set at initialization). Otherwise, the crew must be replaced. A new crew is selected and restarts the cycle with home crew rest. The program will continue to select crews until it succeeds in getting the plane off.

b. Enroute Events

At the time of launch from any base, the simulator ascertains the normal leg time and a random factor, computes actual arrival time, and sets a reminder to act again on the mission at that time, or in some cases an hour earlier as we shall see. If the next base is not a staging point, on arrival time the program must check to see if the crew is qualified to fly the following leg. The program must verify that each member would not exceed his flying time limits on such leg. It must make sure that the maximum duty day will not be violated. If the flying time limits or the maximum duty day would be exceeded or if three consecutive days of maximum duty have transpired, the crew is placed on enroute crew rest and the plane is placed on enroute layover until the crew is qualified to fly again. Upon completion of crew rest, they are allowed time to report to the flight line and pre-flight begins for the following leg.

If the next base is a staging point a notice will be given one hour before arrival to alert a stage crew to report for pre-flight at arrival time. The arriving crew goes directly to crew rest after completing their post-flight inspection. They will be placed in the staging pool as soon as they finish crew rest. Crews in this staging pool rotate on a first-in, first-out basis.

Upon arrival the program samples its ground time distribution and sets a departure time for itself. If, as the clock runs, the actual ground time exceeds the scheduled time, the outgoing crew is placed on ramp status. If the ramp time is less than six hours (or as initially specified), the mission continues. Otherwise, the crew is replaced. In general, this does not cause a large enroute layover of the plane at a stage point since the replacement crew normally has already taken the enroute crew rest. That crew must, however, perform its own pre-flight inspection of the aircraft. The program's strategy calls for alerting this replacement crew one hour before the unscheduled maintenance is completed. If less than one hour of unscheduled maintenance remains, the plane is placed on enroute layover status until the crew arrives. The crew is allowed the usual hour to report to base operations.

c. Post-mission Events

As soon as the plane arrives at home base, the program checks to see if minor or major maintenance is required. If maintenance is not needed, post-flight inspection is performed. The plane is then returned to the available pool. Those planes requiring maintenance are returned to the available pool as soon as maintenance is completed. The crew members are granted free time based on the amount of time they have been away from home this mission. As each crewman completes his free time, the program will check to see if a period of adjustable duties is pending. Before placing him on such duties, the program insures that they will not cause cancellation of a mission. Eventually, the men are returned to the available pool, completing the cycle of all possible statuses.

Output Measurements (what can be analyzed afterwards)

During the course of the simulation run, a log is maintained of every change in status of every man and every aircraft. This log is recorded on a history tape in which each transaction consists of one change in status of one individual or plane. This history tape can then be used as source data for various summaries to describe what happened and for analyses that compare this run to other runs. Just as in operational activities themselves, there are many different variables that might be summarized and many different ways of summarizing each one. Since this analysis phase is not truly an integral part of the simulation itself we will merely mention at this point a listing of some variables which have been computed, by names which are self-explanatory. They are separated roughly into four categories: those pertinent to individuals, those pertinent to crews, those pertinent to planes, and those pertinent to the system as a whole.

a. Individual measures.

1. Total time spent in adjustable duties by month or by crew position or both.
2. Total time spent in unscheduled leave by month or by crew position or both.
3. Total time spent in fixed duties by month or by crew position or both.
4. Total time spent as free time by month or by crew position or both.
5. Average length of the free time period by month or by crew position or both.
6. Distribution of length of free time period.
7. Average time away from home by month.

8. Distribution of time away from home by month.

9. Average time between missions by month.

10. Distribution of time between missions by month.

11. Average flying hours per person by month.

12. Distribution of flying hours per person by month.

b. Crew Measures

1. Average length and distribution of pre-flight time. Note that this measure merely confirms that the program is actually sampling the distribution which was initially submitted to it.

2. Average and distribution of in-flight times per leg. This is really more of a system measure which is pertinent to the formula approach to crew management. Out of the simulation we get merely what we initially put in, within random variation.

3. Time spent in post-flight duties.

4. Time spent in enroute crew rest.

5. Time spent in enroute waiting. This is an important indicator of the staging effect. It will appear in many of the analyses.

c. Plane Measures

1. Number and length of home layovers.

2. Time spent in pre-flight.

3. Time in-flight by leg and by mission and by month.

4. Number and length of enroute plane layovers. This measure reflects those periods when crews were an impediment to the system.

5. Time in minor and major maintenance.

6. Utilization rate. This figure, the number of air hours per day per plane, is the most often used management measure of plane availability and workload. It is used in planning, but appears here as utilization rate actually achieved per period.

d. System Measures

1. Missions scheduled.

2. Missions rescheduled.

3. Missions cancelled.

4. Mission departures.

5. Mission arrivals.

6. Mission legs delayed, and distribution of delay time.

7. Rescheduling delay, total and average time.

8. Work-month-hours. This is a measure of personnel availability or utilization most often used in a manpower management and planning context to the extent that it reflects an individual's contribution to the Air Force out of his thirty days. It is also an individual measure, of interest both as a squadron average and as a distribution from minimum to maximum.

A TYPICAL SIMULATED MISSION

We will illustrate the workings of the simulation program by describing an actual run we made early in our investigation. Even at that elementary stage of evolution we were able to set many characteristics comparable to actual operations while keeping the run as a whole reasonably simple and uncomplicated in order to assure ourselves that it was indeed doing what it was supposed to do. Actually, in any use of a simulation program it is important to keep the values and procedures used as simple as possible consistent with the objective of the investigation. This keeps the computer processing within the spatial limitations of the computer, minimizes the amount of time involved in running on the computer, and insures that any differences arising in the result under different runs are indeed due to the factor or factors being investigated rather than to some undesired cause arising out of some complicated attempt to reproduce reality in its fullest detail.

For this run, we chose the number of planes assigned to the squadron (input variable 13) to be 16, as a number corresponding to reality that we felt could be handled within the computer storage available to us. We chose to define for this run just one kind of mission (input variables 8 and 9); it was called mission type 10, specifying the use of a "basic" crew. We described the route as being outbound from Charleston to

Dover to Elmendorf to Yokota to Clark and in-bound from Clark to Kadena to Elmendorf to Dover to Charleston. Stage pools were defined in each direction at Elmendorf with a pool at Yokota and another at Kadena. Scheduled ground time was set at two hours at each base. The eight legs were given scheduled air times respectively of 1.5, 8.0, 8.5, 4.5, 2.0, 8.5, 7.3, and 1.3 hours. Anyone who has been over this route will recognize that the numbers used do not precisely describe the actuality but that they contain enough truth to represent the operational essence of this route. The maximum duty day was held at 16 hours throughout this run. As a result of this definition of mission type 10, aircraft and crew will fly an average of 41.6 hours per mission, and the air time per leg will average 5.2 hours. Finally, we furnished as part of the initialization an estimate of 4.667 days away from home for the computer to use in scheduling crews to see if the mission would interfere with some future fixed duty.

Next we decided that when we prepared the exogenous event tape we would schedule two missions per day having a fairly constant spacing and yielding a workload that would average out to 83.2 air hours per day or exactly 5.20 as a planned utilization rate. We chose to make this run with 40 pilots, 40 co-pilots, 40 navigators, 40 flight engineers, and 40 loadmasters, giving an equal number of resources at each crew position (input variable 12) although the program is written to accommodate unequal numbers. This total of 40 crews corresponds to a crew ratio of 2.50 which is consistent with our desire to run at first with simple numbers that would not stress the system. The "basic" crew type (input variable 1) was defined as one pilot, one co-pilot, one navigator, one flight engineer and one loadmaster without distinguishing between individuals qualified as command pilot and co-pilot.

As a policy for awarding free time at the end of each mission (input variable 2) the factor of one-third was established for this run, with an upper limit of three days for any one mission. If a mission could not be launched within 18 hours of its scheduled departure (input variable 3), then the program was instructed to cancel the mission. The limits placed on flying time per individual were for this run (input variable 4) 125 hours in 30 days and 300 hours in 90 days. Required crew rest prior to starting a new work day was set at 12 hours (input variable 5). Maintenance (input variable 6) of 12 hours was required when between 258 and 300 air hours had been flown on that aircraft since the previous minor maintenance; "major" maintenance of 96 hours was required when between 798 and 840 air hours had been flown on that aircraft since the previous major maintenance. One flight crew was placed initially in each of four stage pools (input variable 7) as a simple and efficient use of the rather limited number of crews available.

For this run a distribution of ground time (from crew reporting in to take-off) was inserted (input variable 10) only after scanning the operational experience to that date with the C-141. The simulator was initialized to choose randomly each such ground time before take-off, from a distribution which would guarantee in the long run that roughly

27	percent would be between	1	and	1.6	hours
40	percent would be between	1.6	and	2.4	hours
23	percent would be between	2.4	and	3	hours
4	percent would be between	3	and	4	hours
3	percent would be between	4	and	8	hours
2	percent would be between	8	and	16	hours
1	percent would be between	16	and	24	hours

When one of the longest times occurred, of course, the crew would "burn out" by exceeding its duty day and would be replaced by another crew if available.

To account for variations in air time for a particular leg, due operationally to such things as differences in weather and differences in aircraft, we supplied a factor (input variable 11) having a roughly normal distribution from .90 to 1.10 with mean of 1.00; consequently all air times were within ten percent of the standard (input variable 8) for that leg.

In preparing the tape of exogenous events for this run we created a notice for a mission of type 10 to depart at 0500 and at 1700 each day from home base, the notice to be received 15 hours before scheduled departure. To create the notices of unscheduled leave we first selected a distribution of length of unscheduled leave. For this we specified that 60% of the time it would last 2 days, 25% of the time it would last 4 days, 10% of the time it would last 8 days, and 5% of the time it would last 16 days. This averages out to 3.8 days absence per unscheduled leave. Next we calculated the average number of people to be placed on unscheduled leave each day as follows: divide the number of aircrewmen in the squadron by 114. This number was used in order that each person would average 12 days per year on unscheduled leave, accounting for one day per month in the work-month-hour computation. For our run with 200 people in the squadron this average comes to 1.7544. Then, for each day of the simulation, we sampled a Poisson distribution with that mean to determine the number of men to put on unscheduled leave that day. Having the number for the day, we selected the particular individuals at random from among those of the 200 who were not then already on unscheduled leave. For each such man thus selected we sampled the distribution of length of leave to determine how long that individual should stay on unscheduled leave. Finally, we created a notice on the exogenous tape which specified the beginning and the end of each period of unscheduled leave.

In creating notices for fixed duties on the exogenous tape we blocked out two kinds of duty. Each week one pilot, one navigator, one flight engineer, and one loadmaster were selected to serve a period of seven days on fixed duty. This was intended to simulate the aircrew manpower used for scheduling officers. In addition to this, every member of the squadron served one day of fixed duties each month. These notices of 7-day and 1-day fixed duties were placed on the exogenous tape such that at the end of one period of duty the following one would be made known to the program. Thus the program would be able to check ahead and avoid sending a crewman on a mission which would interfere with fixed duties.

Adjustable duties of two kinds were similarly blocked out on the exogenous tape. One kind provided one day per month for each squadron member; the other kind provided a period of four days of duty every two months to simulate those duties that could be performed when each squadron member was not individually needed for a mission flight or its subsequent free time.

The exogenous tape was then compiled in calendar order. The initialization values were fed into the program and the run was begun. Data were allowed to accumulate on the transaction (history) tape until the notices of the passage of simulator time at the console indicated that 180 days had been simulated.

We will now look at two of the output variables computed from that transaction tape, after the run had been completed, in order to analyze what was going on. First, with the exception of the first two 15-day periods, we achieved the planned utilization rate (UR) of 5.2. These first two periods are not representative because we begin at time zero with all planes at home base. While it would have been possible for us to distribute men and planes throughout the system in an attempt to approximate their condition at time zero it was actually easier merely to run the program itself for a while to achieve this distribution. An average UR of 5.19 was actually achieved for the final ten periods. There were no missions cancelled during this run, while nine missions were rescheduled during the six months' simulation.

With an expected utilization ratio of 5.2 for a 30-day month with 16 planes and 40 crews we would anticipate on the average that each crew member would fly 62.4 hours per 30-day month. On the actual run if we ignore the first one month of getting the system running, the average flight time per person for the last five months of the simulation was 62.52 hours. Now we know that in actual practice some people get much more than this and some less. Such variations would be expected to change from month to month, both as to average and as to distribution within the squadron, and they would change even more when various changes are made in workload, resources and policies.

RESULTS

Since the major factors affecting the performance of the system are the workload stresses laid upon it and the resources made available to it, we made several runs to compare the responses of the system to variations in these factors. For a fixed complement of 16 planes the resources made available can be summarized by the number of aircrewmen or crews made available, expressed succinctly as the crew ratio (CR) or number of crews per aircraft. In order to maintain simplicity and assure that any effects observed can be reasonably attributed to the factor being investigated we ran an equal number of loadmasters, flight engineers, navigators, co-pilots and pilots. As three levels of CR encompassing the range of values of operational interest we chose 2.50, 3.44, and 4.38, which are the ratios resulting from assigning 40 crews, 55 crews, and 70 crews respectively to the squadron.

The workload may be expressed in various ways; for a fixed number of planes such as we are using, it may be conveniently expressed by the utilization rate (UR) which is the average number of hours flown per day per aircraft. By choosing to use the single mission type described above in the illustrative run and by calling for 2, 3, and 4 such missions per day respectively, we established UR values of 5.2, 7.8 and 10.4 as planned utilization rates. By making one run at each combination of CR and UR values and holding all other factors as constant as possible, we created the data that would enable a comparison of the combined effects of these two factors. Nine simulation (3 CR's x 3 UR's) are sufficient to analyze the combined effects. We will now look at the results of this set of nine runs.

The primary outcome of significance to the managers of a jet airlift system is the achieved utilization rate. This is shown for each run, along with percentage figures for missions cancelled and missions rescheduled, in Table I.

TABLE I

UR achieved and cancelled/rescheduled missions

		PROGRAMMED UR		
		5.2	7.8	10.4
	UR achieved	5.2	6.8	5.7
CR = 2.50	Percent missions cancelled	NONE	13	45
	Percent missions rescheduled	3	2	2
	UR achieved	5.2	7.8	9.7
CR = 3.44	Percent missions cancelled	NONE	NONE	6
	Percent missions rescheduled	2	4	6
	UR achieved	5.2	7.8	10.2
CR = 4.38	Percent missions cancelled	NONE	NONE	2
	Percent missions rescheduled	2	4	5

The achieved URs are shown graphically in Figure 1. It can be seen that one cell in this nine-cell matrix showed marked system degradation: the combination of crew ratio of 2.50 and UR of 10.4. In three other cells (CR/UR = 2.50/7.8, 3.44/10.4, and 4.38/10.4) there were problems, as evidenced by the number of cancellations which produced a UR well below that programmed. Obviously the system is being stressed at a programmed UR of 10.4: slightly with a crew ratio of 4.38, a little more with a crew ratio of 3.44, and drastically with a crew ratio of 2.50. At the planned UR of 7.8, cancellations drop the achieved UR noticeably for CR = 2.50; for the other two crew ratios there are no cancellations and the UR is being achieved but the number of reschedulings is slightly higher than in the first column, indicating that the system is probably on the threshold of degradation. In the runs with programmed UR of 5.2, that UR was achieved with no cancellations and a minimum of reschedulings although the reschedulings with the smallest crew ratio of 2.50 averaged slightly higher than under the other conditions, suggesting again that this combination of crew ratio and UR might be just under the threshold of degradation. In summary, an adequate crew pool is a necessity if an airlift system is to avoid system degradation. Conversely there is little systemwide gain for an overly generous crew pool. And this simulation tool provides a meaningful method of determining adequacy.

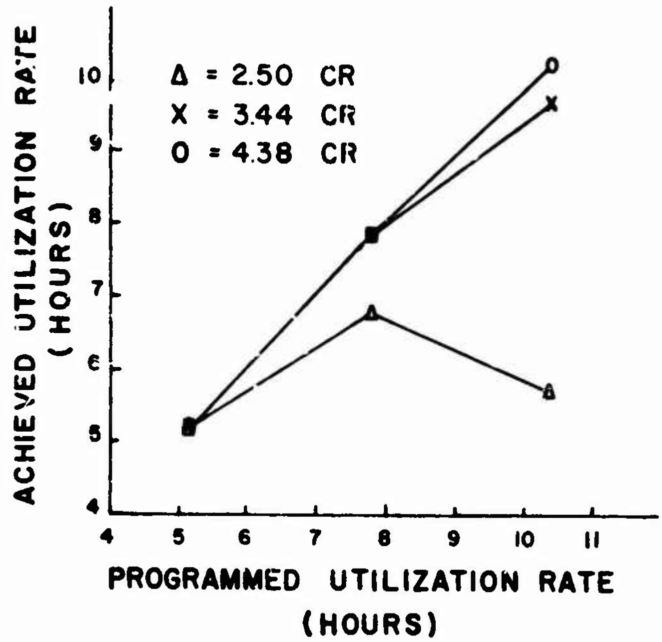


Fig. 1. Achieved vs programmed utilization rates

The effects of increased workload for fixed resources are much as one would have anticipated, with one possible exception. At the crew ratios of 4.38 and 3.44, cancellations appeared indicating stress only at the highest UR while reschedulings gradually increased. At the low CR of 2.50 the system was just able to handle a UR of 5.2, suffered serious degradation at an attempted UR of 7.8, and at an attempted UR of 10.4 achieved a utilization of even less, which may surprise some people at first reading. Basically, in trying to disperse the scarce crew resource throughout the system to accommodate a high rate of flow, the squadron was left with too few at home base to get the missions off the ground and the effects of cancellations kept ballooning. Experienced operational officers will recognize this principle in their own past under highly stressed conditions.

Another major variable used by MAC headquarters in managing its world-wide airlift system is average flying time per month per aircrewman. As in the case of utilization rates there is programmed flying time and achieved flying time. Average achieved flying time per 30-day period for each of the nine runs is displayed in Table 11. As anticipated the flying times increased for fixed CR as UR is increased, except that in attempting to exceed a UR of 7.8 with a CR of 2.50 the average flying time is actually decreased: a drop-off under stress. If the squadron is manned at an effective crew ratio of 3.44, it is able to handle a workload between utilization ratios of 8 and 10 with an average monthly flying time for the squadron of between 70 and 85 hours. Individual times will vary widely from person to person and from month to month in a way that can be deduced from the simulation.

TABLE 11

Average flying time per month (in hours)

	PROGRAMMED UTILIZATION RATE		
Crew ratio	5.2	7.8	10.4
2.50	62	81	69
3.44	45	68	85
4.38	36	53	70

There is another measure which is pertinent to the question of whether manpower is interfering with proper utilization of the system. This is the number of times that a plane is delayed in taking off due to lack of a legal crew. We express it as a percentage of total number of legs flown and display the results in Table 111. These figures indicate that there is little difference whether a squadron is manned at a CR level of 2.50 or 3.44 since the percentage of delays due to lack of crew seems to be essentially the same on each line. Notice, however, that this is a percentage of legs actually flown, so what these figures say is that once a mission is launched on its way we have plenty of stage crews to keep it going. The stress of higher workloads in this set of nine runs was felt in rescheduling and cancellation but not in delays down the line.

TABLE III

Plane delays due to lack of crew
(Percent of take-offs thus delayed)

Crew ratio	PROGRAMMED UTILIZATION RATE 5.2	7.8	10.4
2.50	8	4	3
3.44	8	5	4
4.38	8	5	4

Enroute waiting is a measure which our simulations have demonstrated is very important in indicating whether or not the system is wasting its manpower unnecessarily in the stage pools, as well as a measure which reflects one source of lowered crew morale when it becomes excessively high. One period of enroute waiting begins at the time the crew completes the required crew rest after the previous leg, and extends until the crew begins its travel to the briefing office for the following flight. The average such waiting time in hours is displayed for each of the nine runs in Table IV.

TABLE IV

Average number of hours of enroute waiting at each staging point after completion of required crew rest

Crew ratio	PROGRAMMED UTILIZATION RATE 5.2	7.8	10.4
2.50	11	14	29
3.44	11	11	12
4.38	11	11	11

Here we see the results of a constant stage policy in all but three combinations—these being three of the four where system degradation has already been noted. When the system begins to break down, crews start spending more time in stage pools, are thereby less available for further assignment at home base, thereby creating further cancellations and extending the staging periods of those crews that are sent out, in a vicious circle. The amount and direction of degredation shown by this variable are consistent with that noted in the achieved UR and in the cancellation and rescheduling measures reviewed earlier.

We now come to a set of measures connected even more directly with crew management and crew welfare. The first is the elapsed time between two departures from home base by the same individual, measured in days. The average such time between departures for the whole squadron during the run is displayed in Table V. As one would anticipate, the time between departures increases with increased crew ratio. Similarly the time between departures falls off as the UR increases, except that when the system gets into stress and missions begin to get cancelled the time between departures levels off and even rises, yielding another indicator of break point.

TABLE V

Time between departures (in days)

Crew ratio	PROGRAMMED UTILIZATION RATE 5.2	7.8	10.4
2.50	20	15	18
3.44	28	18	15
4.38	35	23	18

The measure "time at home" is one of the more important areas of concern to aircrewmen. This measure refers to time at home base between missions, measured in days. The squadron average for the time at home between missions for each of the nine runs is presented in Table VI. For a fixed crew ratio, the time at home drops off with increasing utilization, apparently to a bottom or minimum value. Unless the system is unduly stressed, the time at home naturally rises with increasing crew ratio.

TABLE VI

Time at home between missions (in days)

Crew ratio	PROGRAMMED UTILIZATION RATE 5.2	7.8	10.4
2.50	10	5	5
3.44	18	9	5
4.38	25	14	8

The complementary measure, time away from home per mission, again shows (see Table VII) that one affect of degraded system performance is to lengthen the crew time away from home, and the amount of lengthening appears to reflect the amount of stress. Note the slight rise in the top center and right center values as well as the almost 50% increase in the upper right corner over the stable unstressed values.

TABLE VII

Time away from home, per mission (in days)

Crew ratio	PROGRAMMED UTILIZATION RATE 5.2	7.8	10.4
2.50	9.5	10.3	13.5
3.44	9.5	9.4	9.8
4.38	9.5	9.4	9.5

The measure which combines many of the former measures of keen interest both to management and to personnel alike is the work-month-hour measure, which is used especially in manpower planning to describe the number of hours per month that an aircrewman is at work--on duty that is. Let us compute, from the simulated operations, a figure as close as possible to actual work-month-hours achieved. We will do this by adding together in-flight time, pre-flight time, ramp time (crews ready but waiting for plane to be ready), adjustable duties, fixed duties, unscheduled leave, and enroute waiting. Thus, while on a mission, that time spent in required crew rest would not be counted toward the work-month-hour measure but all other time would be. In order to be comparable with other uses of this measure, it seemed necessary for us to make some adjustments. Three of the measures included above would count, for example, an entire day of being sick as 24 hours against the work-month-hour, whereas the proper definition allows only for eight hours of each work day that would be missed due to illness. Similarly, figures for leave are customarily visualized in terms of eight hours per day while they were accumulated at 24 hours per day on our transaction tape. For that reason, in accumulating our work-month-hour achieved figures, we accumulated the unscheduled leave, the adjustable duties and the fixed duties and divided the total by three before incorporating these three components into the work-month-hour total. All other components mentioned above were added in full. The average work-month-hour thus computed is displayed for each of the nine runs in Table VIII. In the low-CR/high-UR combination the personnel were being very highly stressed over an extended six-month period even though, as other measures showed, the production of the system was falling off. The simulation at present does not have any psychological or physiological factors built into it; in actual operations further degradation of the system would be anticipated due to the loss of efficiency by aircrew personnel.

TABLE VIII

Work-month-hours at 1/30 staging policy

Crew ratio	PROGRAMMED UTILIZATION RATE 5.2	7.8	10.4
2.50	210	286	354
3.44	162	220	278
4.38	136	182	227

VERIFICATION OF THE SIMULATION PROGRAM

Although we have already seen that we can depend upon the simulation to function according to the policy rules and distributions which we give it, we should compare simulation results against observed operational results to assure ourselves that there is a modicum of reality in the functioning of our simulation program before we start to draw any further conclusions about operational policies from our simulations. To do this let us turn to operational data which summarized the experience for pilots of the 9th and 75th Military Airlift Squadrons over a six-month period from April to September 1967.

The utilization ratio they experienced was 7.98 hours per day; we will therefore use our results at the programmed UR of 7.8 hours per day. The study reports an effective crew ratio of 2.76 for pilots available and trained, which is 27.66 percent of the way from a CR of 2.50 to a CR of 3.44. We will therefore interpolate between our results for those two crew ratios in order to get comparative figures. We notice further that during the period of the study the number of crews in stage was not exactly that used in the simulations previously described (1 in 30). The average for the three stage points of Yokota, Clark and Wake was reported to be 1 in 40. We have found the staging to have a decided effect on the results, so we must somehow account for this difference in staging. This we can do by using an additional set of runs we made holding stage policy constant at 1 in 45 but with all other conditions the same as before. An interpolation one-fourth of the way from 1/45 to 1/30 between the two numbers achieved for the programmed UR of 7.8 will then produce a simulation figure to be compared to the operational figure.

Let us first look at the work-month-hours, which measure was the principal one used in the operational summary. The total work-month reported for the pilots in these squadrons during this period was 231.8 hours.

On the bottom line of Table IX our simulated work-month-hours for a CR of 2.50 is shown as 241 when adjusted for a 1/40 staging policy. Similarly, simulated results would be a work-month of 185 hours for a CR of 3.44 when adjusted for a 1/40 staging policy. Interpolating between these we calculate a simulated work-month-hour of 225 to be compared with the observed figure of 231.8 hours. This is an exceptionally good corroboration when we consider that the observed figures are averaged over two different squadrons using a varying number of stage crews and are based on overall averages of many kinds of real conditions. Futhermore, our simulated figure could stand to be adjusted slightly upward for the 7% by which the observed utilization ratio of 7.98 lies above 7.8 in the direction of 10.4. We will not attempt to calculate this last adjustment since it would be expected to be rather fine anyway. We will examine, however, the several work elements which go into the total work-month to assure ourselves that the consistency of simulation with reality carries through with each of these work elements.

TABLE IX

Comparison of work-month-hour elements against operational results

		CR = 2.50		CR = 2.76			CR = 3.44	
	1/45	1/40	1/30	1/40		1/45	1/40	1/30
	Simulated	Interpolated	Simulated	Calc.	Oper.	Simulated	Interpolated	Simulated
AF Nonavailable Time	29.4	29.3	29.1	29.4	33.0	29.8	29.5	28.7
Ground Training Sq. Addl. Duties Sq. Alpha Alert	13.4	13.7	14.6	13.2	14.0	11.8	12.0	12.5
Enroute Alert	52.3	70.3	124.2	63.3	66.1	33.3	45.1	80.5
Flying Hours	90.3	88.0	81.1	82.4	81.9	67.8	67.9	68.1
Pre-flight Hours	40.4	39.5	36.7	37.0	36.8	30.6	30.5	30.4
Total work-month	225.8	240.8	285.7	225.3	231.8	173.3	185.0	220.2

Col (2) is interpolated between Col (1) and Col (3)
Col (7) is interpolated between Col (6) and Col (8)
Col (4) is interpolated between Col (2) and Col (7)

In the body of Table IX we present the comparison of the work-month-hour elements interpolated from the simulations against those gathered operationally. The first, third, sixth, and eighth columns in this table are actual results from runs previously mentioned. The second column is interpolated between the first and third columns to adjust for 1/40 stage policy. Similarly, the seventh column is interpolated between the sixth and eighth columns to adjust for 1/40 stage policy. Finally, we interpolate again between the second and seventh columns to adjust for the crew ratio 2.76 and place this result in the fourth column alongside the operational results in the fifth column.

Except for two items these overall averages are only tenths apart. In the case of the Air Force nonavailable time we put into the simulation about 32 hours but got back only 29.4 hours. This is probably due to our handling of crew members who "fell sick" while away from home base on a mission. To avoid having to provide the mechanism for switching the membership of crews while out on a mission, and excusing ourselves on the basis that an aircrew member only slightly ill would prefer to continue with his crew if at all able to function, we simply ignored that sick leave which fell during an away-from-home period, thereby reducing the total time spent in Air Force nonavailable time. The only other element differing by more than half an hour is enroute alert, which differs by less than three hours. When we consider that both of the figures being compared were obtained by rather broad interpolations and composite calculations, and when we consider how sensitive this figure is to stage policy, it is an extremely good fit.

Although only a single comparison of simulation results with data from actual operations is obviously not adequate to prove the realism and validity of the simulation, the results in this case are very encouraging.

CALCULATING A CREW RATIO

We have shown how the influence of crew ratio and utilization ratio on each of several variables could be studied by presenting the results of nine simulation runs in a square array for each variable. Before trying to build a composite picture of the UR/CR effect let us make use of the techniqus presented in an earlier report which releted the outcomes of e simulation in terms of the ratio CR over UR. While in the former report we explored several approaches to this retio, we will here confine ourselves to programmed CR/UR. One of the principal reasons for settling on the use of programmed UR rather than achieved UR is thet the ratio then can be used easily for planning purposes since the programmed UR is known et planning time. While either UR over CR or CR over UR could be used to tell the same information we choose the latter, primarily because a frection less than one will not be confused with one of the ratios UR and CR. For quick reference we display this retio in Table X for each of the CR/UR combinations discussed thus far. We will now show how the verious previously described output measures of the simuletions vary when releted to the CR/UR ratio and then proceed to use this information in discussing the selection of an optimum crew ratio.

TABLE X

Values of the retio CR/UR

Crew ratio	PROGRAMMED UTILIZATION RATE 5.2	7.8	10.4
2.50	.48	.32	.24
3.44	.66	.44	.33
4.38	.84	.56	.42

First, let us consider the principal evidence of system stress--the inability to move the planned loed. One wey to displey this inability is by releting the echieved UR to the programmed UR, which we express as e percent decrease in programmed UR. Precticelly identicel inferences can be drawn here by studying the percentege of mission cancellations, which yields simpler figures. In operetionel situetions theee might not meesure the same thing beceuse there might be e tendency to cancel the longer missions in fevor of the shorter, or some such bias, but in our simuletions the fect that ell missions were alike permits us to use the percentege cancelletions to study this effect. In Figure 2 we have plotted the percentege of missions cancelled egainst the CR/UR retio without maintaining the structure of the simuletion matrices as we have been doing up to this point. In addition to the originally described nine runs having stege policy 1/30 we have displeyed an additional set of six runs having stege policy 1/60. Note that for both policies there is e sherp break around ratio .33, signs of low level system stress between .33 and .47, with zero stress ebove that figure. Although the degree of stress represented by percent cancelletion is less for the staging retio 1/60, the regions of values of CR/UR in which stress occurs seem to be essentielly the same regerdless of steging retio.

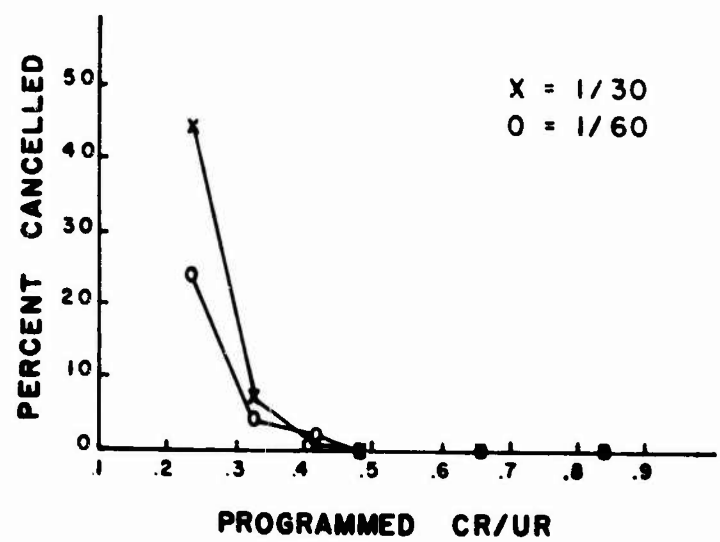


Fig. 2 Percent of missions cancelled vs CR/UR retio

Now let us turn to the question of selecting a crew retio. We have seen how, even efter making other runs for other constant stage policies, our results still consistently reinforce our previous finding that stress conditions do not seem to exist above a CR/UR ratio of about .47, that there is gradual system degradation for CR/UR ratios down to about .33, and that there is a marked rapid degradation below that. We are led to the existence of a zone of CR/UR ratios between .35 and .47 wherein a) the system is at the point of mild stress degradation implying full use of facilities and b) aircrews are functioning at a full but not overly stressed work level. The logic of our previous work for choosing a CR still stands:

1) determine the highest UR that may be programmed any length of time. If a UR is maintained long enough to be recorded in the reporting statistics then presumably it will have been maintained long enough for its stress effects to occur.

2) choose a core crew ratio which must absolutely be above the low end of the zone for this UR in order to avoid serious system breakdowns due to lack of crews.

3) preferably choose a core crew ratio at or above the middle of the slight-degradation zone for this UR.

To be precise consider a systems planner who wished to fly a routine program of utilization ratio 6.5 while maintaining the capability to fly a demand program of utilization ratio 8.0 for extended periods without creating frequent system crises. We will need to use the information in Table XI. Manning at a crew ratio of at least 2.80 is imperative; this would yield a CR/UR ratio of .43 for the routine program--cancellations running routinely around one percent. In the extended periods of UR = 8.0, CR/UR would be .35 with cancellations running to five percent of programmed workload. This represents continual evidence of marginal operations and the ever-present possibility of a serious inability to perform the mission. To accommodate the extended periods of UR = 8.0, a crew ratio of at least 3.44 is advisable. At this level routine operations would be at CR/UR = .53 yielding no cancellations so that the programmed workload is being accommodated without overstressing or under-utilizing crews. Extended periods of UR = 8.0 would yield around one percent cancellations so that the system, including crews, is slightly stressed but still has a reserve margin for peak short-term demands without serious failure.

TABLE XI

Core crew ratios related to CR/UR ratio

PROGRAMMED UR

	5.0	5.25	5.5	5.75	6.0	6.25	6.5	6.75	7.0	7.25	7.5	7.75	8.0
.33	1.65	1.73	1.82	1.90	1.98	2.06	2.14	2.23	2.31	2.39	2.48	2.56	2.64
.35	1.75	1.84	1.92	2.01	2.10	2.19	2.28	2.36	2.45	2.54	2.62	2.71	2.80
.37	1.85	1.94	2.04	2.13	2.22	2.31	2.40	2.50	2.59	2.68	2.78	2.67	2.96
.39	1.95	2.05	2.14	2.24	2.34	2.44	2.54	2.63	2.73	2.83	2.92	3.02	3.12
.41	2.05	2.15	2.26	2.36	2.46	2.56	2.66	2.77	2.87	2.97	3.08	3.18	3.28
.43	2.15	2.26	2.36	2.47	2.58	2.69	2.80	2.90	3.01	3.12	3.22	3.33	3.44
.45	2.25	2.36	2.48	2.59	2.70	2.81	2.92	3.04	3.15	3.26	3.38	3.49	3.60
.47	2.35	2.47	2.58	2.70	2.82	2.94	3.06	3.17	3.29	3.41	3.52	3.64	3.76
.49	2.45	2.57	2.70	2.82	2.94	3.06	3.18	3.31	3.43	3.55	3.68	3.80	3.92
.51	2.55	2.68	2.80	2.93	3.06	3.19	3.32	3.44	3.57	3.70	3.82	3.95	4.08
.53	2.65	2.78	2.92	3.05	3.18	3.31	3.44	3.58	3.71	3.84	3.98	4.11	4.24

SUMMARY

We have described a crew-oriented computer simulation program which yields data on system-wide operation in formats which facilitate management decisions on manning, crew welfare, tolerable workload, and mission effectiveness. Typical data were presented. Verification against operational data was reported. The use of this technique to derive a crew ratio suitable for a wide range of operational demands illustrates but one application of simulation approaches. Both simple and radical departures from existing and aircrew management policies can be explored at low cost.

DISCUSSION

HOLDEN

Most of this work was done on computer simulations and I presume that you made some attempt to verify or get field data which would allow you to recheck some of the assumptions?

HARTMAN

We have three validations. When we built our programme we had a sample system in being and we put in some of the variables. We then ran a prolonged simulation to see if we could regenerate the original data. Further at the time of the Vietnam war we reported that MAC could not sustain an 8 hour per day utilisation rate. At that time MAC had been at the 8 hour per day utilisation rate for about 9 or 10 months, but they found very shortly that it was necessary to cut back because the system as a whole was in difficulty. Finally in our most recent work we have been flying many wars inside a computer with various profiles. We flew the 31 day Israeli episode resupply operation.

von WERNER

Did you simulate the unstable condition of the system and can you make some comment on how long the system takes to stablise. Are there favourable or unfavourable conditions to achieve this?

HARTMAN

We did not intentionally simulate the unstable condition as it occurs spontaneously as a result of negative interactions when the workload is high. For an airlift the answer is about 20-25 days, but from this time the operation begins to degrade.

A SIMULATOR STUDY TO INVESTIGATE HUMAN OPERATOR WORKLOAD

by

P.H. Wewerinke
J. Smit

National Aerospace Laboratory NLR
Anthony Fokkerweg 2
Amsterdam-1017
The Netherlands

SUMMARY

This paper presents the results of an exploratory experiment which was conducted to investigate human response characteristics in control situations of widely varying difficulty. The experiment was aimed at a better understanding of the human operator limitations in terms of control effort as included in the optimal control model.

Based on the experimental results a control effort index is presented. The "predicted" control effort correlates excellently with subjective ratings and seems to have a useful generality.

1 INTRODUCTION

The ever increasing complexity of man-machine systems necessitates a more complete description of the human capabilities and limitations.
For the modelling of complex real-life problems the systems-approach has proven to be feasible. The description of the human operator as a subsystem involves three sets of variables:

- . <u>input variables:</u> environmental conditions, task variables
- . <u>internal variables:</u> the state of the organism, physiological and psychological states and potentialities
- . <u>output variables:</u> overt behaviour, psychological and psychophysiological phenomena.

In the field of manual control several "black box" models of the human operator have been developed (e.g. quasi-linear pilot models, optimal control models). They are based on output-input relationships and offer the possibility to "predict" the control behaviour and system performance for a range of control tasks. This approach has resulted in a good knowledge of the control characteristics of man.

For the allocation of functions in complex man-machine systems (e.g. manual vs. automatic) and for the specifications of the man-machine interactions (display-control compatibility) it is also necessary to have an insight in the interactions between the input variables and the internal variables. The demand of the task (in terms of attention required) combined with the environmental conditions can be labelled as workload.
The effects of this load on the human has to be assessed via measurements of output variables such as performance, quality of information processing, and psychophysiological parameters.

Given a certain criterium (either a well-defined, external criterium or a subjective, internal criterium) the operator has to exert a certain amount of effort to meet this criterium when the task becomes more demanding the operator can maintained his performance level at the cost of more effort investment. This can be conceptualized in a limited capacity model of human information processing (Ref. 4). The demand of the task will be reflected in the portion of the capacity involved in accomplishing the task.
Secondary task methods have been used to determine the capacity which was not consumed by the primary task.
Another method is to measure the variations in some physiological parameters as a function of a variation in the demand of the task. These methods have not yet resulted in models of human operator workload which can generally be applied.

Even for manual control problems in a restricted sense (pure perceptual-motor tasks without monitoring and planning) the concept of effort so far has hardly been incorporated in the design and development of man-machine systems.

In the next section a human operator model is discussed in terms of both human performance (optimal control model) and control effort. An exploratory experimental program is described which was conducted to build and support the given control effort index.

2 OPTIMAL CONTROL MODEL OF HUMAN OPERATOR PERFORMANCE AND EFFORT

2.1 Optimal control model.

In this Chapter the principal features of the optimal control model of the human as a feedback controller developed by Kleinman et al (Ref. 1) are briefly discussed. The model is based on optimization and estimation theory and can be used for multivariable linear control systems. This approach is more appropriate for the analysis of complex man-machine problems (high workload, multiloop control situation) than the conventional servo-systems approach which relies heavily on judgements concerning the closed loop system structure.

The model is based on the assumption that the well-motivated, well-trained human operator behaves in a near optimal manner subject to his inherent limitations, and constraints and his control task. The representation of the human limitations, as included in the model, is discussed next. Paragraph 2.1.2. contains a system and task description.

2.1.1 Human limitations.

The psychophysiological limitations of the human operator which are included in the model are:

1. A lumped equivalent perceptual time delay τ, representing the various internal time delays associated with visual, central processing and neuromotor pathways.
2. Controller remnant which is taken to be the component of human response that is unpredictable in other than a statistical sense. The various sources of human randomness are represented by errors in observing system outputs and executing control inputs by including observation noise and motor noise in the model. Observation noise might represent the effects of random perturbations in human response characteristics, time variations in response parameters, random errors in perceiving display variables. In general it will also depend on type, quality and form of the displayed information. Motor noise can be associated with randomness in executing the intended control movements and (or) the fact that the human does not have perfect knowledge of the control inputs.
3. Neuromotor dynamics represented by a first order lag indicating that the human operator is not able or willing to make rapid control movements.

The foregoing can be summarized referring to Figure 1. It is assumed that the human perceives a delayed, noisy replica of the system output which is processed by an equalization network. This network represents the means by which the human attempts to optimize his control strategy. The commanded control ($\underline{u}_c$) perturbed by motor noise is operated on by the neuromotor dynamics to provide the control motion $\underline{u}$.

2.1.2 System and task description.

The human operator's basic task is to control, in some way, a dynamic system. Both the characteristics of this system (dynamic characteristics, disturbances involved, displayed variables, etc.) and the criteria which the human operator is instructed (able or willing) to optimize, determine system performance and the load imposed on the human operator to achieve this performance.

The mathematical representation of the task environment, the human operator's control objectives and his limitations is (among others) given in Reference 1.

2.2. Control effort.

2.2.1 Background.

Levison et al (Ref. 2) define a workload index as the fraction of the controller's capacity that is required to perform a given task to some specified level of performance. The concept of this index is based on the assumption that the human operator possesses a fixed amount of (channel) capacity to be shared among these tasks. The generality of this model is limited because the workload index is highly task-dependent and "calibration" experiments are necessary to determine the human operator's "full" capacity in a given control situation. Next the fraction of this "total" capacity can be determined in multiple-task situations. As such, the workload index serves only as a means for comparing the relative load imposed on the human operator by various tasks. The next paragraph contains a brief discussion of an experimental program which was conducted to provide data for a variety of control situations to build and validate an absolute workload model.

2.2.2 Experiments.

In order to include all possibly important characteristics of pilot behaviour in control situations of (our) interest, a variety of single axis control tasks were performed by four well-trained, highly motivated subjects (experienced fighter pilots). The choice of the tasks was determined by two characteristics: pilot adaptation to the task (position-, rate-, and acceleration control) and the sensitivity of task performance to the effort exerted by the human operator (level of instability). The task was to regulate against a disturbance input.

The experimental results (Ref. 3) were presented in terms of scores of system parameters of interest, frequency domain measures (describing function and remnant data), and subjective information obtained by means of rating scales. Model parameters (time delay, lag factor, observation noise, and motor noise) were obtained by fitting the experimental data.

The results have shown that

- for a variety of control situations the available measures of human operator behaviour can accurately be duplicated by the optimal control model by primarily a variation in the pertinent noise ratios
- the values of the noise ratios are task dependent
- the effort involved in achieving the pertinent noise ratios is clearly determined by the task.

Based on foregoing observations a control effort index is proposed in the next paragraph consistent with some notions of attention and effort of the experimental psychology (Ref. 4).

2.2.3 Control effort index.

Human operator behaviour is partly determined by mechanisms that control the choice of stimuli, by which is meant both selectively attending to some stimuli in preference to others and investing more or less attention per source of information. This can be identified with voluntary attention reflecting that the subject attends to the stimuli because of their relevance for performing the task and not only because of their arousal function.

Also involuntary attention has to be included in a concept of effort. This can be related to the level of arousal (Ref. 5) and is largely dictated by the properties of the stimulus situation. Processing novel and surprising stimuli involves more effort than in the situation of more familiar stimuli.

Based on the foregoing notions a control effort index is defined which is highly compatible with the optimal control model. The former aspect (voluntary attention) is incorporated in terms of signal to noise ratio of the various sources of information. This can be identified with the amount of attention as indicated in Reference 2. The latter aspect (involuntary attention) is included in the control effort metric in terms of sensitivity of task performance (mean-squared error or error rate) to the momentary attention paid by the subject as indicated by the pertinent signal to noise ratios.

In formula

$$E = \frac{S}{f}$$

with

$$S = \frac{\partial \sigma_x^2 \, (dB)}{\partial f \, (dB)}$$

where σ_x^2 is the variance of quantity x which is taken as the relevant performance index

In control situations where more than one source of information is of interest (position information, rate information, etc.) the workload index is defined as the sum of the separate portions.

The straightforwardly computed control effort is compared with the subjective ratings obtained during the experiments for various control tasks. The result is shown in Figure 2. The relationship between the subjective ratings and the control effort index is surprisingly linear.

It is interesting to consider the relationship between system performance and effort. For a typical control task (rate control) this is shown in Figure 3. The trade-off between performance and effort is as might be expected:
exerting little effort in the low effort region gives a relatively large improvement in performance. Beyond a certain level of system performance the pay-off for more effort is very small. One would expect the optimum trade-off at the "knee" of the curve. The experimental results confirm this.

3 CONCLUDING REMARKS

The experimental results give an additional demonstration of the usefulness of the optimal control model as a tool to describe human operator behaviour in a variety of control situations.

The control effort index suggested in this paper correlates very well with subjective ratings. By means of trade-off studies between system performance and control effort the effect of task variables of interest in new control situations can be studied.

More control situations (also multiple control tasks) have to be considered to be able to extend the generality of the proposed control effort index.

4 REFERENCES

1	Kleinman, D.L. Baron, S.	Manned Vehicle System Analysis by Means of Modern Control Theory. NASA CR-1753, June 1971.
2	Levison, W.H. Elkind, J.I. and Ward, J.L.	Studies of Multivariable Manual Control Systems: A Model for Task Interference. NASA CR-1746, May 1971.
3	Wewerinke, P.H.	Human Operator Workload for Various Control Situations. Paper presented at the 10th Annual Conference on Manual Control, Dayton USA, April 1974.
4	Kahneman, D.	Attention and Effort. Prentice-Hall Inc. 1973.
5	Berlyne, D.E.	Conflict, Arousal and Curiosity. McGraw-Hill, New York, 1960.

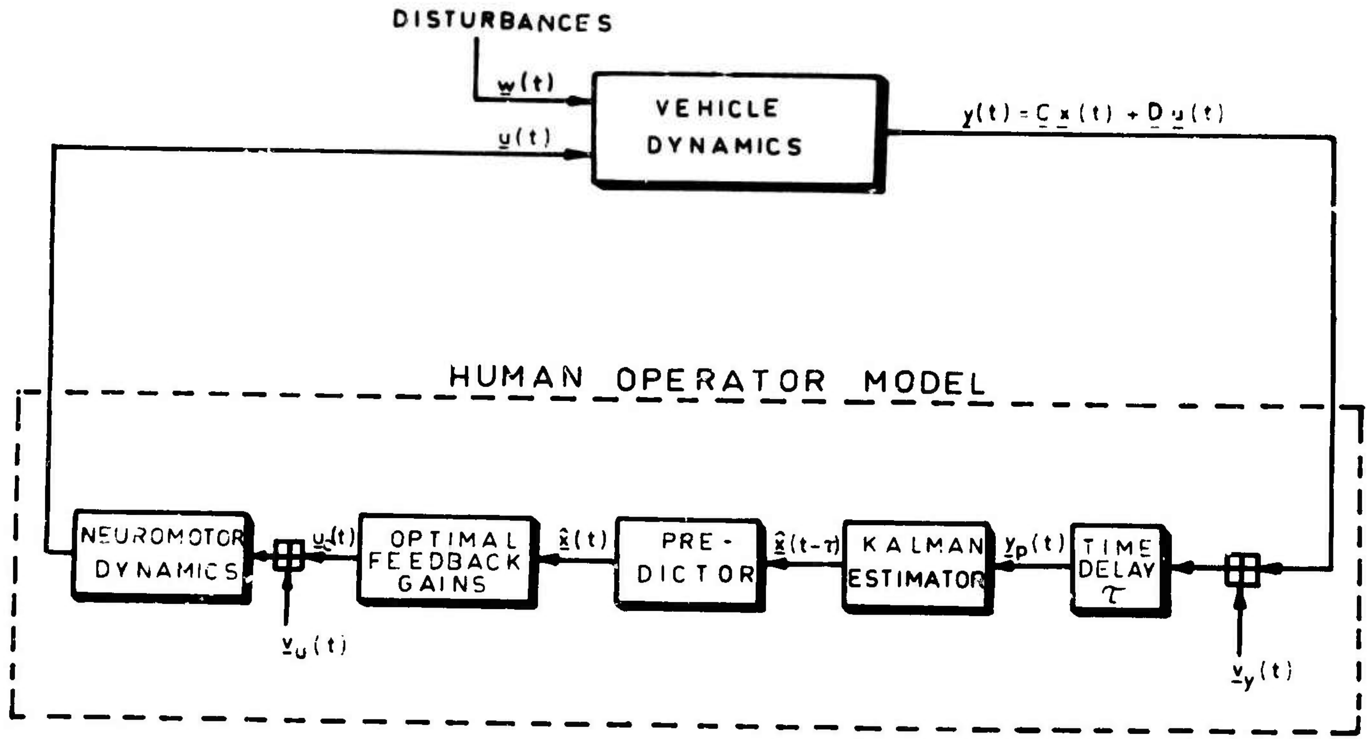


FIG. 1 BLOCK DIAGRAM OF THE HUMAN OPERATOR MODEL

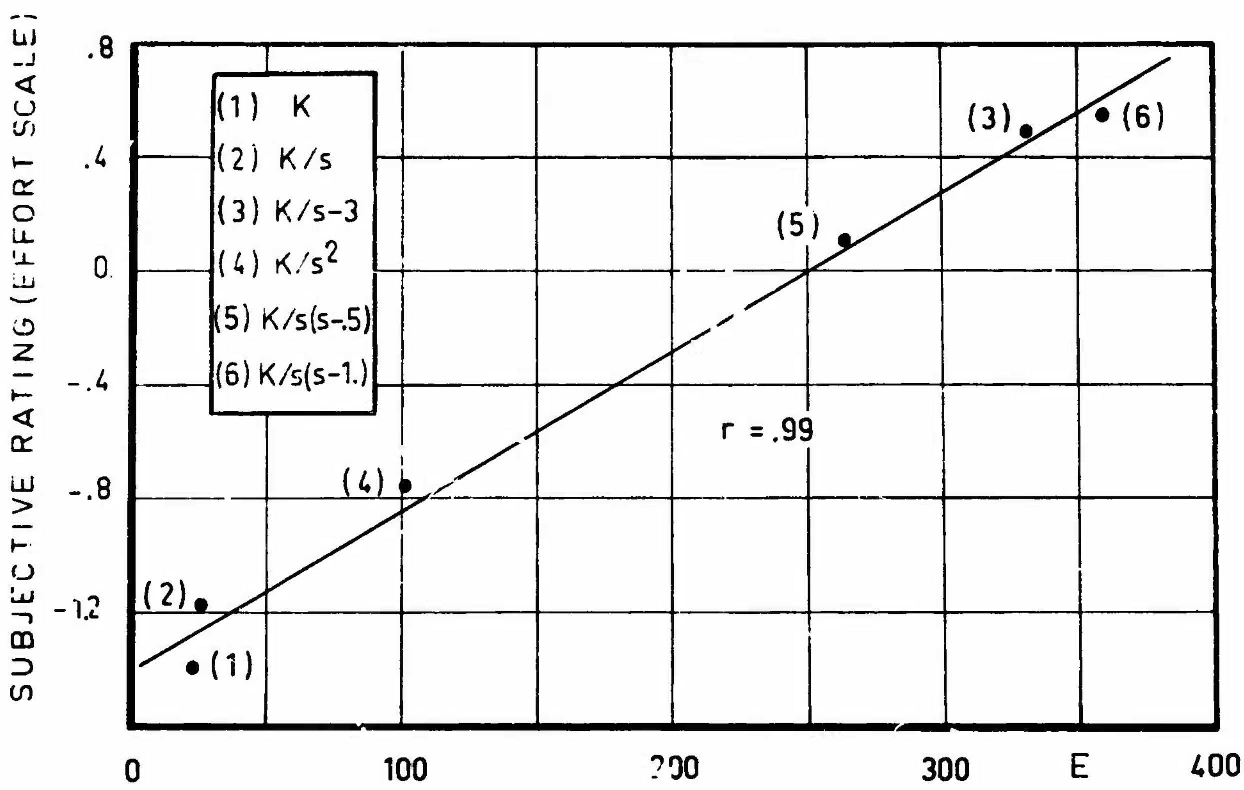


FIG. 2 THE RELATIONSHIP BETWEEN SUBJECTIVE RATINGS AND COMPUTED EFFORT.

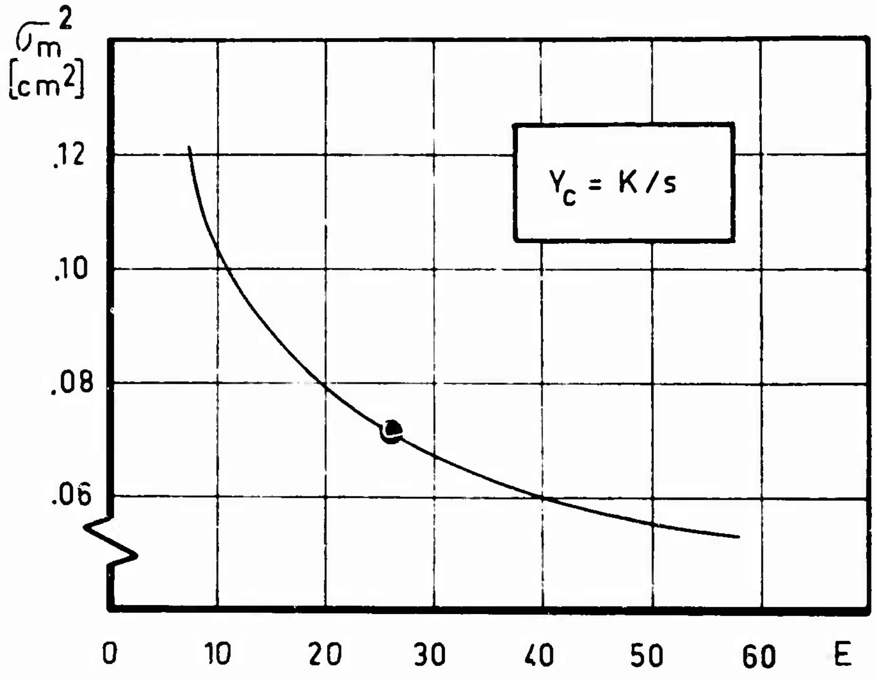


FIG. 3 THE RELATIONSHIP BETWEEN SYSTEM PERFORMANCE AND EFFORT.
$Y_c = K/s$

DISCUSSION

HOLDEN Across a number of differant controlled alements you could presumably identify the noise disturbance required to come out with an equivalent workload measure for ell the differant controlled elements. If that is true would you expect tha pilot subjective ras.onses to be the same in tarms of evaluating workload?

WEWERINKE That's right. That would be a possibility. We would have tha opportunity to very ell kinds of tasks and environmentel veriables and see how they effect the workload.

NICHOLSON In your model do you assume that a subject assesses workload by tha same technique under high workload conditions as undar low workload conditions?

WEWERINKE We obtain retings by prasenting the scelas and esking tha subjects to inaicete the level of effort they had to expend. We wera not awara of tha mental processes involved so I cannot make e very scientific statement about that.

LABORATORY RESEARCH INTO HUMAN INFORMATION PROCESSING

by

Jo H F Huddleston
Human Factors Group
Royal Aircraft Establishment
FARNBOROUGH HAMPSHIRE UK

WORKLOAD MEASURES

An hour or two in a technical library soon shows that we have begun the development of three main families of measurement in the (acute) workload area. Glibly, we can measure how well a pilot is meeting some set of physical criteria (objective), we can assess autonomic and related indices of task arousal/involvement (physiological), and we can ask the pilot what he thinks (subjective).

"Subjective" data can never be complete and accurate. I cannot describe fully and precisely how I hold my pen or read my own English in public, nor can I agree to relinquish part of my esteem and salary because myriad other people might have the same skill. What holds for me holds equally for the piloting fraternity, and it would be churlish and naive to expect them to be able to describe their personal, occupational or social condition coldly and in detail. Given a little intuitive statistics, however, one can see that asking, say, 12 pilots whether subtask X is 0.4 or 2.3 times more demanding of visual attention, say, than subtask Y might be a defensible but piecemeal way to proceed. Inviting a test pilot to confess he cannot meet a challenge, or a line pilot to admit he is overpaid, is asking for more than verbal trouble, however.

"Objective" data are not so much complex, in the amusing intellectual sense, as voluminous, in the boring sense. After some systematic cataloguing of what processes the pilot is controlling, there follows a listing of likely measures. Thus height keeping, for example, can be represented meanly as a needle-chasing task, and the deviation between actual and demanded needle positions can be spot-checked or time-averaged in some acceptable way, and control utilization similarly recorded. In this sense, cockpit skills are normally exercised in a context which is rich in data, to the point of embarrassment and occasional despair. Nevertheless, one can impose additional (unnatural) tasks if one thinks their measurement is simpler. Herein lies the "secondary task" philosophy which, to my mind, has been more clever than effective.

"Physiological" measures seem to spring in and out of fashion, perhaps as researchers exhaust one line and attack another. They are always disturbingly indirect in their implications, however. Popular examples will serve to argue this: heart rate is controlled by more physiological etc subsystems than task-induced stress, and eye-point-of-regard may often be to the featureless geometric midpoint of several sites to which one is paying attention.

Currently, there is much hope (and little evidence) that a triumvirate of measures will be helpful. Thus we talk of measuring subjective, objective and physiological activities concurrently, without really attending to the logic and arithmetic of the case. For performance measures cannot be understood without asking the pilot what his work criteria were; physiological indicants of involvement/arousal similarly demand that the pilot name what was concerning him; pilot opinion neither reflects overt activity precisely nor remains uncorrupted by personal predilections. It is improbable that there will be an easy agreement between such measurement families, or hence any general solution. It is more likely that the finance expended in this search for the psychophysiologist's stone will be more than the differential risk capital of the specific avionics or procedures options at issue.

LABORATORY PSYCHOLOGY

Crudely speaking, Human Factors seems to attract two kinds of workers. There appear to be those who sympathise with humanity and scorn what has been called "the counting fever", and those who believe in the precise, step-by-step application of scientific method and who speak disrespectfully of any other "wishy-washy nonsense". Though this is a caricature, it will do to make the point that the latter "type" of expert finds he can survive in aerospace circles, and the former feels that entry is forbidden him.

Now the psychophysiology of intelligent, well-motivated adult males, riding the sharp ends of metal tubes through weather, is likely to be highly complex. Not surprisingly, those Human Factors experts who volunteer for this area tend to go for experimentation which cannot be attacked by their scientific and engineering fellows. Sadly, this gives rise to precise data which lack the vision due to any member of mankind. It describes pilots as needle-chasers, or requires more accurate language from them than can reasonably be expected. At the limit, it yields methodologically unassailable studies of constrained situations which do not encapsulate flying.

As examples, recent literature in experimental psychology can be considered as if it divided into "pure" and "applied". In the former area, great store is placed on studies of spoken messages to different ears or of simple pictures flashed at different eyes. To be scientifically controlled, such studies have to be uncomfortably simple, so that a great gulf is fixed between arithmetical results and aeronautical conclusions (if any). On the bright side, we are winning a painful understanding of how man defends himself against an onslaught of imprecise data and reaches those brave 2-choice decisions necessary for survival. In the "applied" area, tracking tasks, time-shared with some visual or auditory monitoring, still hold pride of place. Unfortunately, manual tracking is becoming obsolete faster than we are understanding it. At best, it never achieved more than a very demeaning model of man.

WHAT SHALL WE DO?

I shall ignore contradiction, and offer four points of argument. First, we could give up; secondly, we could continue as now; thirdly, we could bring more people (more money) to bear; fourthly, we could spend longer wondering where to place our few bets.

It will be no surprise that the first two courses of action do not have my sympathy. As to giving up altogether, the penalties of failure are frankly too great, especially when one notes how quickly the market and technology can change. As to continuing current practice, it seems to me that this would be to expect diminishing returns from a sustained investment. For the nicely manageable areas of human factors were marked out for treatment before about the mid 1960s. Even the remaining ergonomics issues turn out to have vast queries embedded in an apparently simple structure. To give an example, the straight question: "How do we design a warning panel?" turns into the asking: "What are brains good at and what are computers good at, and how do we mate the two?".

Option 3, to buy more expert activity, seems justifiable to a degree. Aircraft losses, for example, may have a greater than 50% "human error" attribution of one kind or another. This rather suggests that we are cleverer at providing tools than we are at understanding workers, and argues that some cut-back in tooling costs may allow rather better labour relations to be nourished.

Option 4, to consider our bets more carefully, appeals to my taste for intellectual gambling. To expand the argument a little, let us pretend that aircrew workload has the following components:

1 A knowledge of goal-tactic-strategy.

2 Lookout.

3 Flight path control.

4 Navigation.

5 Communications.

6 Engineering systems management.

7 Social.

The first item includes an appreciation of a vast range of ambitions, from "Continue the human race - by earning cash and social respect" to "Hope the weather broadcast is repeated - I'm too busy changing frequencies now". It is a uniquely human ability (at the moment) to switch from macro to micro and back, or choose to linger at intermediate levels, in a hierarchy of ambitions like this.

Item number two involves human vision, either through cockpit transparencies or of electronically-generated display screens. As with item one, this skill is (at the moment) a uniquely human attribute. The welter of detail which has been added to the literature in the last century serves to mask our undisturbed ignorance of what visual perception is really about. As a gross oversimplification, each individual spends his initial 5 years moving round and feeling those objects of which his eyes demand investigation, while the sensors and processor result from millions of years of evolutionary R & D. We shall not mimic this skill yet awhile.

Items 3, 4, 5 and 6 constitute the main stamping ground for engineering and human factors effort in the aerospace field. In the main, they cover problems to do with quite well-understood physical phenomena, such as systems dependent on electrons moving through metal with, as the final stage, some aircrew interfacing. At the limit, it would be possible to automate the pilot out of such systems (though it would be neither cheap nor popular). A blank CRT screen, say, would mean all pre-programmed aspects of altitude, course, systems and messages were being tolerably dealt with. A message on the screen might be of the kind "Unable to accept 35,000 ft: can offer 37 or 39" or "No 3 engine has been shut down" or "Non-routine message awaits re hydraulic maintenance". It is relatively easy to argue in these areas; it is relatively easy to see what R & D deserves being undertaken; it is relatively easy to conceptualize the appropriate man-machine complementarity.

Item number 7, like some aspects of items 1 and 5, is not a person-automaton interface but a person-to-person happening. For one of the few ways we have of estimating the integrity and effectiveness of a skilled human performer is to ask an equally skilled performer to do the estimating. For example, peer ratings, as a method in the work-study area, are used to bridge the alleged gap between "objective" and "subjective" measures as defined earlier. At its slightest, the technique appears in job-unrelated conversation between cockpit crew members, and serves for example to assure pilot and navigator that each is in a presentable condition today. It goes without saying that this kind of activity is not to be automated foreseeably.

Hence "pilot workload" can be expressed as the coping with two main clusters of problem. One cluster is equipment-based and theoretically amenable to physical modelling, even replacement. The other cluster is emphatically to do with humanity (and man-made tools would perhaps never be welcome here); to do with those elusive personal gambles on which life is felt to depend. At any rate, pilot workload might arguably be the maintenance of a mental model encompassing these two kinds of time-varying input; the continued provision of a mental solution for two sources of uncertain problem. If that is so, then the most we can do at the moment is respectable transfer-function work; we can vary inputs, monitor outputs, and debate black box contents with a rare freedom.

So far, the nicely-manageable areas (in the "physics" problem cluster) have attracted most finance, so we have an understanding, defensible to engineers, of the engineering-like inputs and outputs. The shifty, social areas (in the "humanities" problem cluster) continue to frighten us. We would like to acknowledge their importance, but they will not yet fit the narrow mould of scientific method, and we may feel they never will.

Would anyone care to bet, with me, that avionics-based workload is not so much crucial as nicely manageable?

EVALUATING MEASURES OF WORKLOAD USING A F'IGHT SIMULATOR

by

J. M. Rolfe
J. W. Chappelow
R. L. Evans
S. J. E. Lindsay

RAF Institute of Aviation Medicine, Farnborough, Hants

and

A. C. Browning

Royal Aircraft Establishment, Farnborough, Hants

SUMMARY

Any attempt to measure workload must be based upon a satisfactory definition of the nature of the workload that is being investigated. Workload can be considered as having task oriented and subject oriented elements; a variety of measures can be applied to the assessment of either or both of these elements. Moreover, the experimental evidence available suggests that if a comprehensive analysis of workload is to be achieved a range of measures must be employed.

The paper describes an experiment in which a flight instrument trainer, resembling a twin jet communications aircraft, was used to evaluate questionnaire, performance and activity analysis measures of pilot workload. Attempts were made to distinguish between the physical, perceptual and mental components of workload. For this purpose three flight plans were devised, of approximately equal duration, differing markedly with respect to the three above components. Six professional pilots flew each flight plan and after landing completed questionnaires to assess the workload levels and the task content. During the flights video recordings were made of the pilot's manual and communication activity. Performance during ILS approaches immediately before and after the experimental flight plans was also measured.

From these measures it was possible to obtain significantly different results relating to the different flight plans. These results were capable of distinguishing between the three components of workload represented in the flights.

INTRODUCTION

Workload may be considered as having task oriented and operator oriented elements. The task oriented element is synonymous with task difficulty and is related to the nature of the task and the quality of the equipment provided for the performance of the task. The operator oriented element takes into account not so much what the task is but what the task costs the individual in terms of his effort, fatigue etc. A variety of measures can be applied to either or both of these elements. Moreover, the experimental evidence to date suggests that if a meaningful analysis of workload is to be achieved a range of measures must be employed (1, 2). In the experiment described in this paper three particular types of measure were employed and assessed namely, direct performance measures relating to the quality of aircraft handling, observation of the pilots' physical activity and subjective assessments by subjects and observers. Activity recording and subjective assessment were chosen because recent field investigations have indicated that they appear to offer a viable means of gathering descriptive information (3, 4). However, it was argued by the authors that the accuracy of the information obtained from these measures required verification. It was therefore decided to undertake an assessment of the information provided by these in a situation where the nature of the flight task could be controlled and systematically varied to emphasise different aspects of the pilot's work.

METHOD

The investigation was undertaken using the Institute's flight simulator research facility (Fig. 1). The device, a procedures trainer, is representative of the Hawker Siddeley HS125, a two crew twin jet aircraft. The cockpit unit is mounted on a motion system having pitch and roll axes. There is no representation of the external visual world. Inside the cockpit both pilot and co-pilot are provided with a full main instrument panel (Fig. 2). The first pilot's panel includes a Collins FD 108 integrated flight system, an Omni Bearing Selector (OBS) and Distance Measuring Equipment (DME). All primary and secondary flight controls are represented and an auto-pilot is provided.

Three experimental flight plans were prepared for the evaluation by a military pilot in current flying practice. They were designed so that as far as possible they involved grossly different types of activity, namely physical, perceptual or mental workload.

The physical flight plan (Flight A)

The physical flight plan was designed as an 'air test' in severe turbulence. Air traffic control directed the aircraft's position and progress throughout the flight, minimising the pilot's navigation, timing and procedural workload. The pilot was required to carry out manoeuvres involving rapid changes of heading, height and aircraft configuration under normal and asymmetric power. The auto-pilot was not used.

The perceptual flight plan (Flight B)

This flight plan required the pilot to carry out an unorthodox and complex approach procedure involving the use of most of the navigation instruments to maintain position accurately whilst turning continuously in a large arc. Throughout the flight the RT background was noisy and the pilot was required to make many mode changes in his flight system. His attention was directed more than normally to the engine instruments by two occurrences of low oil pressure.

The mental flight plan (Flight C)

This flight took the form of a timed navigation exercise followed by an accurately timed 'fly past' over a known datum point. The pilot did not know the wind velocity and the 'overhead time' was passed to him approximately ten minutes before he was due to overfly the datum point. Physical and perceptual workload was reduced as much as possible by allowing the pilot to use the auto-pilot.

All three flights started at the same airfield. After take-off the aircraft turned right and executed an ILS approach and roller landing at another airfield close by. After clearing the circuit from the second airfield the pilot carried out one of the three flight plans described above. At the end of the experimental phase, the pilot returned to a point common to all three flights, and carried out an ILS approach and landing at the same airfield at which the ILS was executed at the beginning of each flight.

Prior to the commencement of the evaluation all subjects undertook a programme of familiarisation in the simulator. These were designed to acquaint the subject with the handling characteristics of the simulator and the operation of the Collins FD 108 Flight System. The order in which subjects flew the different experimental flight plans was balanced. Before each flight the subject received a detailed briefing. The flight was then flown in the form described above. In each of the experimental flights the subject was accompanied by a co-pilot who not only performed the specified duties of the co-pilot but also observed and assessed the subject's performance.

The three categories of measurement; performance, observation and subjective assessment were used in the following ways:-

Performance Measures

During the ILS approaches at the beginning and end of each experimental sortie recordings of glide slope and localiser deviation, air speed, aileron and elevator activity were made on FM tape and processed using a DEC PDP 12 digital computer. Recordings were made between the outer and middle markers on the approach (between 5.2 and 1.25 miles) and digitized at a rate of 5 samples per second. Localiser and glideslope data were reduced to mean deviation score per 3 second epoch and the mean modulus deviation was used for analysis. Air speed was reduced to a score of the mean air speed per 3 second epoch and the standard deviation of these data was used in the analysis. For elevator and aileron activity all changes in control position of greater than 2 arbitrary units (stick position being calibrated between 0 and 100 representing full displacement from extreme to extreme) and lasting for longer than 0.3 of a second were defined as a movement. The analysis was based on the number of such movements recorded during the period of the approach examined.

Observational Measures

To record the pilot's activity a TV camera with a wide angle lens was mounted on the roof of the simulator and positioned so that it was possible to record movements of the subjects' hands anywhere in the cockpit (Fig. 3). The television picture together with a sound recording of all communications messages between pilot and co-pilot, and pilot and ground control were recorded using a Sony Video-Recording System. The method of analysis of the pilot's hand movement activity was by means of a keyboard, the buttons of which represented the cockpit controls. An observer watching the video play back of pilot activity shadowed the movements of the pilot's hands with two fingers moving over the keyboard (Fig. 4). The keyboard was linked to the PDP 12 digital computer which calculated the total time for which each control was used and the distribution of times spent on each control. Communication messages were analysed in terms of the number and length of the messages to and from the pilot.

Subjective Measures

The pilot subjects completed three questionnaires at the end of each flight. The first questionnaire was task oriented and listed a range of activities. The subject was asked to indicate the extent to which each was present in the flight just completed by making a mark on a 10 cm line, the extremes of which were defined as 'totally absent' and 'extremely prominent'. The activities assessed were derived from two sources:

a. From the literature on the taxonomy of task description (5).

b. From a preliminary study in which aircrew were asked to define the activities which they considered were part of aircraft operations.

Fourteen categories were used for assessment. In the main these were sub-divided into three sections - gathering information, processing information and taking action. In addition the subjects were asked to make an assessment of the opportunity provided by the flight for relaxing. A full list of the activities is included in Annex A.

The second questionnaire was operator oriented, seeking to obtain the pilot's assessment of the nature and extent of the mental, perceptual and physical workload he experienced during the experimental flights. In all a total of 21 individual scales were presented to the subject asking for ratings of factors such as the extent of mental and physical effort expended and the extent to which he considered the flight to be complex, exhausting, difficult and demanding. Again 10 cm rating scales were employed. In this case the extremes were denoted by 'not/none at all' and 'a very great deal'. The scales employed are listed in Annex A.

The third questionnaire combined ratings from the task oriented and subject oriented areas. Four categories of activity were chosen - 'monitoring', 'calculating and estimating information', 'logging information' and 'making adjustments to the aircraft's primary controls'. For each of these categories of activity the subject was asked to make the following assessments:

a. How busy were you in this activity throughout the flight?

b. In this activity how much effort did you spend searching (looking, listening and watching) for information?

c. In this activity how much mental effort did you spend throughout the flight?

d. How much physical effort did you spend altogether on this activity?

The subject made his response on a similar scale to that employed in the second questionnaire.

Observer Ratings

The observer who acted as co-pilot for the experimental flights rated the subject's performance in respect of the following:

Checks.
Take-off and climb.
Aircraft handling.
Instrument flying.
RT procedure.
Procedural flying.
Operating the flight system.
Approaches.
Airmanship.

For each of the above categories the observer assigned a mark out of a hundred with a mark of less than forty indicating failure to perform the activity satisfactorily.

Subjects

Six subjects participated in the evaluation. All were volunteers drawn from pilots working or residing in the immediate area of the Institute of Aviation Medicine. Five of the subjects were military pilots and one was a commercial pilot flying with a national airline. The subjects had an average total flying hours of 2,500 hours with a range from 1,500 - 3,500 hours.

RESULTS

Before presenting the results obtained from the several measures of workload, it is in order to discuss the performance of the subjects as assessed by the co-pilot observer. Table 1 shows the mean rating given to each subject in each of the three flights.

TABLE 1

MEAN OBSERVER RATINGS OF PERFORMANCE

	FLIGHT		
Subject	A	B	C
1	79	59	69
2	78	74	77
3	58	67	66
4	77	71	70
5	82	71	85
6	80	71	70

All the subjects were rated as performing above the 'fail' mark. There was no significant difference between the flights in respect of mean performance ratings, or the ratings on individual scales (Friedman two-way analyses of variance). On only one scale was a rating of less than 40% achieved - by subject 3, who was given 38% for take-off and climb in Flight A. Thus, in the opinion of the observer, the general level of performance was satisfactory and did not vary significantly from flight to flight.

Performance Measures

The analysis of the ILS recordings was in three parts. The first concerned differences between the three sorties which were compared in respect of the changes in performance and control activity between the first and second approaches. Table 2 presents a summary of these changes.

TABLE 2

MEAN PERFORMANCE AND CONTROL ACTIVITY SCORES

	FLIGHT								
	A		B		C		MEAN		Percentage change between flights 1 and 2
Approach	1	2	1	2	1	2	1	2	
Measure									
Localiser deviation (arbitrary units)	26.7	24.7	29.9	24.1	28.6	30.9	28.4	26.6	- 6%
Glideslope deviation (arbitrary units)	11.6	14.4	18.2	15.1	15.5	17.3	15.0	15.6	+ 6%
Airspeed variability (arbitrary units)	5.2	6.9	4.5	5.8	6.2	5.4	5.3	6.0	+ 17%
No of elevator movements	61	74	68	80	54	68	61	74	+ 22%
No of aileron movements	69	73	59	69	71	82	66	75	+ 12%

No statistically significant differences between the three flights were found for any of the above variables (Friedman two-way analyses of variance).

The second comparison was of performance on the first and second approaches irrespective of the intervening sortie. No significant differences were found in relation to localiser and glideslope deviation or airspeed variability. However, the number of control movements, both aileron and elevator, was greater on the second approacb ($p < .05$ in both cases: Wilcoxon matched pairs signed ranks tests). Expressed as percentages, on the second approach there was a 22% increase in the mean number of elevator movements and a 12% increase in the mean number of aileron movements.

In order to discover if any other changes in control activity had taken place the data were re-examined to obtain measures of the mean amplitudes and durations of both elevator and aileron movements. The means obtained are shown in Table 3.

TABLE 3

MEAN NUMBER, AMPLITUDE AND DURATION OF CONTROL MOVEMENTS

		APPROACH	
		1	2
Elevator	Mean No of movements	61	74
	Mean amplitude (arbitrary units)	4.84	5.08
	Mean Movement time (secs)	0.72	0.72
Aileron	Mean No of movements	66	75
	Mean amplitude (arbitrary units	10.83	11.16
	Mean Movement time (secs)	0.87	0.84

In both cases no difference in amplitude or duration between the first and second approach could be found. This result led to the conclusion that while the number of movements increased during the second approach the type of movements made remained unchanged. Also, if the finding was taken in conjunction with the performance measures, there was no indication of any change in performance accompanying the increase in control movements.

The third part of the analysis looked for differences between the subjects. Using Kruskal Wallis one-way analysis of variance tests these were found to be present for three of the variables, glideslope deviation ($p < .02$), elevator activity ($p < .02$) and aileron activity ($p < .01$). Product moment correlation coefficients were calculated between these three items. The values obtained are shown in table 4 below.

TABLE 4

PRODUCT MOMENT CORRELATIONS BETWEEN CONTROL ACTIVITIES AND GLIDESLOPE DEVIATION

	Elevator Activities	Glideslope Deviation
Aileron Activity	+ .70 ***	- .63 **
Elevator Activity		- .43 *

* = p < .05
** = p < .01
*** = p < .001
N = 24

These relationships suggest a consistency in the pattern of control activities and that greater amounts of control activities are associated with smaller deviations from the glideslope.

Consistency of control activity style was further examined by two comparisons of the number of aileron and elevator movements made by the subjects on the first and second approaches.

The relationship between the number of elevator movements on the two approaches was positive ($p < .05$) and produced the linear regression equation:

$$Y = 29.82 + 0.70428\,X$$

where X = the number of movements made on the first approach and
Y = the number of movements made on the second approach.

The relationship for aileron movements was also positive ($p < .01$) with the linear regression equation:

$$Y = 13.87 + 0.95241\,X$$

Observational Measures

Hand activity was examined in terms of the cumulative time spent on the various controls in the cockpit. Table 5 shows the percentage time spent with the hands on the control column, on other controls and off the controls.

TABLE 5

HAND ACTIVITY: PERCENTAGE TIME ON CONTROL COLUMN ON OTHER CONTROLS AND OFF THE CONTROLS

		FLIGHT		
		A	B	C
Left Hand	Off controls	5	4	28
	Control column	94	95	9
	Other controls	1	1	63
Right Hand	Off controls	11	17.5	40
	Control column	69	55	2
	Other controls	20	27.5	58

This analysis showed differences in hand activity between the flights. It also showed differences between left and right hand activity. In Flights A and B the left hand was on the control column almost all the time (94% in Flight A and 95% in Flight B). In Flight C when the autopilot was engaged the left hand spent much more time on other controls (63%) but also was less active and spent more time off controls.

Compared with the left hand the right hand spent more time on other controls in Flights A and B (20. and 27.5%). It also was off controls more during these flights. In Flight C the right hand was also less active and the time spent on other controls increased.

Table 6 contains a breakdown of the time spent on controls other than the control column.

TABLE 6

HAND ACTIVITY: PERCENTAGE OF TIME ON CONTROLS OTHER THAN CONTROL COLUMN

		FLIGHT		
		A	B	C
Left Hand	Stop watch	-	-	4
	Flight documents	-	-	41
	Flight system	1	1	18
Right Hand	Flight system	1	2	6.5
	Altimeter	-	1	-
	Pitch trim	5.5	8	-
	Throttles	12.5	13	-
	Flap selector	-	1	-
	Radio	-	2	1.5
	Autopilot	-	-	4
	Rudder trim	1	-	-
	Flight documents	-	0.5	34

The differences in the nature of the tasks comprising the three flights are reflected in the use of other controls. In Flight A, containing manoeuvres under assymetric power, the bulk of the activity is related to the use of the throttles and trim controls. Flight B, containing an approach procedure, has activity related to power, trim, radio and flight system. Flight C is different again, with the activity being associated with the use of flight documents, the flight system and autopilot.

A similar method was used to analyse communications activity and the results are shown in Table 7.

TABLE 7

COMMUNICATIONS ACTIVITY; PERCENTAGE TIME IN EACH CATEGORY

	FLIGHT		
	A	B	C
Pilot speaking	9	11.5	9
Others speaking to pilot	28	20	5
Silence	63	68.5	86

Flights A and B again contain more activity than Flight C and with the pilot being spoken to more than speaking.

Subjective Measures

In the case of the task oriented questionnaire significant differences between the flights were found in seven of the fourteen rating scales. Table 8 summarises these data.

The subjects reported less activity in gathering information from maps, documents and displays other than the main flight instruments in Flight A than in B or C. Flight C was reported to involve more calculation, estimation and logging of information than Flight A. Flight A was rated as requiring more control movements, both large and small, than Flight C, and Flight B required more small control adjustments than C. Operating the autopilot was rated as more prominent in Flight C than in A and B.

TABLE 8

TASK ORIENTED QUESTIONNAIRE: MEAN RATINGS AND SIGNIFICANT DIFFERENCES

	FLIGHT				
	A	B	C	Significance Level	Indication
Rating Scale					
Obtaining information from subsidiary instruments	27	61	60	*	A < B & C
Retrieving information from charts and documents	9	38	44	**	A < B & C
Calculating and estimating information	29	39	66	**	A < C
Making large movements to a/c primary controls	64	48	28	*	A > C
Logging information	6	23	43	**	A < B & C
Making small adjustments to a/c primary controls	68	66	45	*	A & B > C
Operating the autopilot	3	0	49	**	A & B < C

* $p < 0.05$ ** $p < 0.01$

Five of the twenty-one scales in the operator oriented questionnaire distinguished between the flights. Table 9 summarises these data.

TABLE 9

OPERATOR ORIENTED QUESTIONNAIRE: MEAN RATINGS AND SIGNIFICANT DIFFERENCES

	FLIGHT				
	A	B	C	Significance Level	Indication
Rating Scale					
Physical effort	52	29	22	*	A > B > C
Perceptual effort	48	71	80	**	A < B < C
Complexity	31	51	74	**	A < B < C
Exhaustion	49	36	46	*	A > B & C
Confusion	8	46	2	*	A < B

* $p < 0.05$ ** $p < 0.01$

The subjects reported the expenditure of more physical effort in Flight A than in B, and more in B than in C. The greatest amount of perceptual effort (looking, listening, watching for information) was expended in Flight C, the least in A. Flight C was rated more complex than A and B, and Flight A more exhausting than B and C. Flight B was thought more confusing than Flight A.

The third questionnaire showed significant differences between the flights with respect to the distribution of mental effort among four major activities. Table 10 summarises these data.

In Flight A the major part of this effort was devoted to control adjustments. In Flight B it was to monitoring displays. In Flight C it was to calculation and estimation.

DISCUSSION

The three types of measure employed in this investigation all provided information of varying value in differentiating and describing the three sorties in relation to the task and operator oriented elements of workload.

The performance measures recorded at the beginning and end of the experimental sorties provided no information which discriminated between the three sorties. There was the finding that irrespective of the sortie flown the number of control movements made during the second approach and landing was greater than that made during the first. However, no changes in amplitude or duration of movement accompanied the increased number of movements and the approach performance was not different from that produced during

TABLE 10

RATINGS OF MENTAL EFFORT ASSIGNED TO VARIOUS ACTIVITIES

	FLIGHT		
	A	B	C
Activity			
Monitoring (m)	44	58	65
Calculating and estimating (c)	56	48	79
Logging (1)	26	20	46
Making control adjustments (a)	67	31	32
Level of significance	**	**	**
Indication	1 < a & c	m > 1	1 < m & c

the first approach. These results have more value to the description of operator behaviour than to task demand. The measures of performance show individual differences and these are best interpreted as indications of variations in particular 'operator style' which relate not so much to the quality of the performance as to the mode of control behaviour used to achieve the required level of performance. This observation is supported by other studies (6) and the present data are currently being re-examined to look at these individual differences in more detail.

The recording of operator activity generated valuable information and a viable technique for the translation of operator behaviour into an objective and quantifiable record. The major value of the activity measure is seen as contributing to an understanding of the task oriented element of the workload. The results obtained distinguished between the three flights in terms of the amount of activity and the locations at which this activity took place. Similarly the analysis of communications gave an indication of the relative frequency of communication and its direction. The data derived from these measures are clearly more relevant to ergonomic considerations than to workload. The inspection of transitions between controls, the frequency with which controls are used and the length of time spent on particular controls, provided information which can be used to assess the design effectiveness of particular configurations. In addition the technique provides observational measures which allow an assessment to be made of the operator's behaviour when performing checks, procedures and cockpit management generally. Failures in these areas can also be identified which are not revealed through traditional performance measures. For example, one subject performed all his calculations in Flight C aloud. The activity recording revealed an error of 90 seconds in one of his calculations which persisted until the subject, suddenly realising his mistake and its effect upon his management of the flight, retrieved the situation with a drastic change of power setting and configuration. Such errors would normally come to light in the real world when they are catastrophic.

In relation to the operator oriented element of workload the activity analysis provided little information. It indicated what the subject did but not what it cost him to do it. Again there are indications of individual differences between subjects in relation to the frequency of movements, the time spent on particular controls and hand preference. These findings are being investigated as part of the examination of individual differences.

With regard to the operator activity measures it is worth noting that the technique of visual shadowing requires very little training for the observer. An opportunity did arise to allow a comparison of inter-observer variability. Using three different observers analysing the same portion of video record very little inter-observer variability resulted and a consistent picture of the operator's patterns of activity was obtained from all three observers.

Information regarding both the task and operator elements of workload was obtained from the subjective rating scales. Using the operator to undertake his own assessment of task content lead to the identification of a number of important activities. The results were consistent both with the contents of the sortie as initially designed and with the observation of operator activity. It is, however, recognised that the three flights were intentionally designed to be markedly different from one another, and it remains to test the technique in more subtly varying situations.

The operator oriented questionnaire revealed differences in the reactions of the subjects to the three flights. It is of particular interest to note that whilst the scales differentiated between the three flights in terms of the amount of physical and perceptual effort it was not initially possible to discriminate between them in terms of the amount of mental effort each contained. This finding suggests that the subjects recognised that so long as the task demanded their conscious attention there was an element of mental load present to which could be added varying amounts of physical and perceptual load depending upon the extent to which control display configurations and tasks varied. The third questionnaire clarified this situations by identifying that the amounts of mental load subjectively considered to be present, while similar, were attributed to different elements in the task situation. Again, these differences reflected the differing nature of the three sorties. Finally, there is the indication that different descriptions of the subjective feelings of the operator in relation to his experience of task load are differentially associated with the varying sources of load. Thus, for example, the term 'complexity' appears to be associated with mental load, whereas assessments of the feelings of exhaustion experienced by the subjects are more related to physical load.

This initial trial using the Institute's flight simulator facility has provided information which can guide further experimentation and development. The use of and refinement of subjective assessment techniques for examining both the task and operator aspects of workload appears to be worth pursuing. A combination of these measures with objective performance measures and the observation of activity would appear to offer a useful perspective on the global effects of cockpit layout, task content and operator experience of workload and fatigue.

REFERENCES

1. ROHMERT, W. & LAURIG, W. Work measurement. Psychological and physiological techniques for assessing operator and workload. Int. J. Prod. Res. 9, 1971, 157-168.

2. ROLFE, J.M. & LINDSAY S.J. Flight deck environment and pilot work load: biological measures of work load. Applied Ergonomics 4, 1973, 199-206.

3. HOPKINSON, J.A. & SHACKEL, B. Summary report of the flight-deck activity study. Loughborough University of Technology Department of Ergonomics Report. LUTERG 83, 1972.

4. ROLFE, J.M. & CHAPPELOW, J.W. The application of aircrew opinions on cockpit tasks and equipment to flight safety research. AGARD CP-132, 1973.

5. CHRISTENSEN, J.M. & MILLS, R.G. What does the operator do in complex systems? Human Factors 9, 1967, 329-340.

6. HUDDLESTON, H.F. Developing measures to reveal individual styles in the use of an acceleration control. Int. J. of Man-Machine Studies, 4, 1972, 33-44.

ANNEX A

SUBJECTIVE ASSESSMENT SCALES

(1) Activities rated on a scale: Totally absent - Extremely prominent.

Obtaining information from main flight instruments.

Obtaining information from all other instruments, displays and gauges.

Retrieving information from charts and documents.

Obtaining information from aircraft crew and ATC.

Calculating and estimating information.

Choosing between alternatives.

Passing information to crew and ATC.

Making large movements of the aircraft's primary controls.

Logging information.

Making small adjustments and corrections to the aircraft's primary controls.

Operating other controls in the aircraft.

Relaxing.

Operating the autopilot.

Monitoring.

(2) Operator oriented rating scales.

How vigilant were you throughout the flight?

How relaxing was your experience throughout the flight?

How pleasant was the experience throughout?

During the flight how automatic was your behaviour?

How much physical effort did you expend altogether?

Throughout the flight how much effort did you spend in searching (looking, listening and watching) for information?

How tiring was the flight for you?

How uninterrupted was your behaviour throughout the flight?

How active were you throughout the flight?

Were you tense throughout the sortie?

During the flight as a whole how attentive were you?

How busy were you throughout the flight?

How much mental effort did you expend altogether during the sortie?

How demanding was the flight?

Was the sortie difficult on the whole?

Was the flight confusing?

Was the flight boring?

Was the sortie unusual?

How strenuous was the sortie as a whole?

How exhausting was the sortie?

Was the sortie complex?

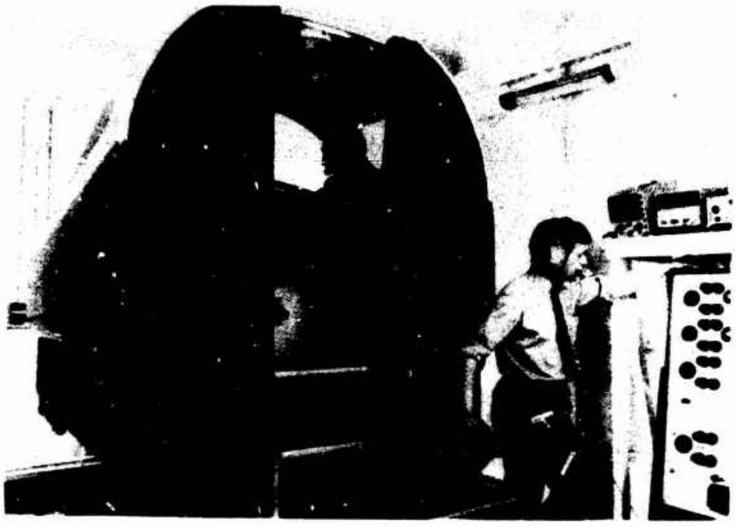


Fig 1 General view of IAM Flight Simulator Facility

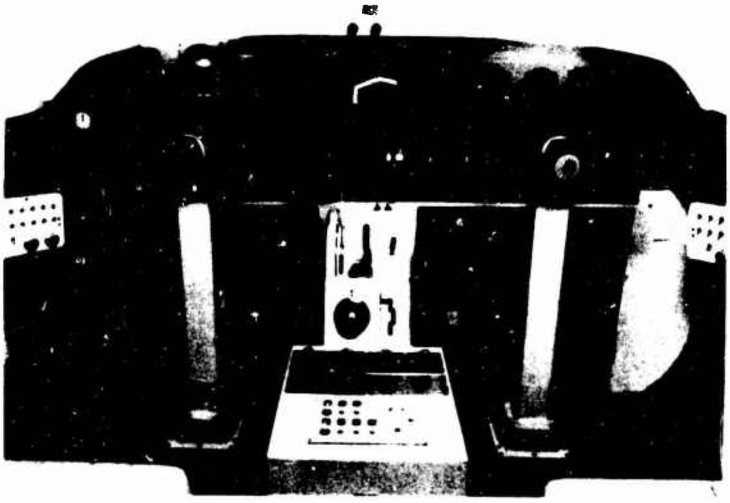


Fig 2 Cockpit layout of simulator

Fig 3 View from video camera position

Fig 4 Observer monitoring pilot's activity using keyboard

DISCUSSION

WILKINSON

Was there any correlation between subjectively rated mental effort and the observer's rating of the level of performance, even if the variations in neither of these was very large?

ROLFE

There were some correlations but others which we expected were absent. The workload may be covert and inherent in some other aspect of the job which was downstream from the final overt activity.

A FLIGHT SIMULATOR STUDY OF MISSILE CONTROL PERFORMANCE AS A FUNCTION OF CONCURRENT WORKLOAD

by

K.G.G. Corkindale
Royal Air Force Institute of Aviation Medicine
Farnborough, Hampshire, England.

SUMMARY

Eight pilote took part in a part-task simulation of the delivery of a stand-off air-to-surface guided weapon. The attack phase of a sortie wae eimulated. This phase lasted some 3 minutes and included a low level run to the weapon release area, weapon release, target detection on the TV monitor display and the aiming of the mieeile at the target. Four levele of workload were etudied. These were:

1. Missile control taske only.

2. Manual control of the eimulator.

3. Missile control tasks plue manual control of the simulator.

4. Missile control tasks plue auto-pilot monitoring.

The resulte showed that:

1. Performance at the miseile control tasks was degraded by increasee in concurrent workload.

2. Manual flight control and auto-pilot monitoring were adversely affected by concurrent missile control taske.

A emall group of non-pilots wae put through the same experimental programme. A eimilar pattern of reeulte to those of the pilot group were obtained but the absolute levels of performance were predictably different.

Eye movement data and subjective data allow the deterioration in two-task performance to be explained. The results of this study confirm laboratory secondary task experimente in that a deterioration in primary task performance ie associated with the occurrence of a secondary task deepite instructions given to the subjecte to maintain the higheet level of performance poeeible on the primary task.

INTRODUCTION

The research flight eimulator offers the systems designer the opportunity to asseee, at an early etage in a development programme, the effects of alternative equipments and working proceduree on aircrew performance. The alternativee that are to be compared may be seen as providing different levele of aircrew workload and the designer'e interest is in how these different workload levele will affect aircrew performance and the implicatione of any changee in performance on system capability.

During studiee of the operator's taske in controlling a etand-off air-to-surface televieion guided weapon (TV ASGW) a flight eimulator trial wae conducted in which the effecte of other aircrew tasks on miesile control performance were examined. Four working conditione were studied, these were:

1. Missile control tasks only. The taske were to monitor the television dieplay for target detection, to align cross-hairs on the target and to home the missile on to the target by continuously tracking the target as it moved across the TV screen.

2. Simulator manual control using a Head-Up Display (HUD). A standard pattern HUD was presented to the pilot who manually flew the flight eimulator according to the HUD demande.

3. Missile control tasks plus HUD manual control. The pilot carried out the TV missile taske (as in 1 above) and also manually flew the simulator (as in 2 above).

4. Missile control taske plue HUD monitoring. The pilot carried out the TV missile taske and was also required to monitor the HUD whilet the eimulator was flown on auto-pilot. As a check on the efficiency of the monitoring, the pilot was required to respond to an infrequent eignal on the HUD by preesing a button mounted on the control column.

These four working conditione allowed two major comparisons to be made. Firstly, the effect on the TV tasks of alternative concurrent workloade. Secondly, the effect on aircraft control of the additional workload caused by the TV taske. As the pilot was required to perform two separate taske, his primary task being control of the aircraft and the secondary task being missile control, the situation can be interpreted as an example of the measurement of mental load by the secondary task technique. Rolfe (1971) has reviewed the extensive work on the secondary task technique. Hie main finding was that deepite the inetructione given to subjecte the presence of a secondary task depressee performance on the primary task, although in some cases changes in performance on the primary task could not be clearly identified due to the use of insensitive measures of primary task performance. It was intended that the present study could provide a check on this general finding when the two tasks were seen by the subjects as having some real-life significance.

Another question of some interest that was examined during the trial was the use of non-pilots as subjects in flight simulator studies. The main experiment used experienced jet pilots as subjects but a second experiment used non-pilots who although familiar with the simulator, through having worked with it as experimental staff for some months, had no aircraft piloting experience.

METHOD

The flight simulator was a single-seat fixed-base cockpit unit provided with a control column, rudder pedals, throttle and intercom system. Conventional head-down instruments and the associated controls were not represented. An outside world view could be provided by means of a slide which gave an indication of ground-horizon-sky visible from the pilot's seat. This slide was controlled by outputs from the simulator computer so that pitch and roll information was available to the pilot. No representation of forward movement was provided. The pilot's primary information source was a HUD which operated in a director mode. For the manual control runs the pilot was required to keep the aircraft symbol centred over the target marker. In the HUD monitoring mode the signal which the pilot was required to detect was the appearance of a small cross on the display.

The TV display was generated on a 200 mm (8 inches) 625 line monitor situated some 750 mm (30 inches) from the eye datum position and some 20° down from the horizontal forward line of sight. The TV target was a small white square which could be made to appear on the TV tube face in any one of twelve pre-set positions. From the time this target was switched on it followed an exponential growth law which approximated to the rate of change of contrast of a target in the real world when being approached at uniform speed. Two cross-hairs working in display x and y were controllable by a simple displacement joystick control mounted just aft of the throttle lever and worked by the pilot's left hand.

The task presented to the subject was a simulation of the attack phase of a stand-off TV ASGW mission. In detail this part-task simulation consisted of the following events:

1. High speed low level run (90 seconds, TV display off).

2. TV weapon release, TV monitor switched on.

3. Weapon run to target area (42 seconds, TV on).

4. TV target bright-up started.

5. Pilot detects target on TV, presses button on control column to call cross-hairs onto TV display.

6. Pilot centres TV cross-hairs on target and presses control column button.

7. Pilot tracks target with cross-hairs as the target is moved across the display for 10 seconds, at the end of which time TV is switched off.

8. High speed low level run (20 seconds).

The average total time for each experimental run was some 3 minutes 10 seconds, the exact time varied slightly due to different target detection times and cross-hair aiming times. This sequence was varied slightly in the case of the simulator manual control runs (working condition 2) as no TV tasks were presented to the subject who was required to track using the HUD for 3 minutes and 10 seconds. Similarly, for the missile control tasks only runs (working condition 1) the initial and final low level runs were shortened to minimise the time the subject sat in the cockpit with nothing to do. In the TV tasks plus HUD monitoring runs (working condition 4) two signals were presented on the HUD, one during the first 90 second period and the other during the target bright-up phase; the exact timings of these two signals were randomised.

The eight pilot subjects were all experienced on high performance jet aircraft, their average age was 30 years and their average total jet hours were 1,600. The four non-pilot subjects were drawn from the staff who maintained and ran the simulator and although they had some simulator flying experience they had no aircraft piloting training or experience.

On arrival at the simulator the subjects received a standard written briefing on the trial and were familiarised with the equipment they were to use. The subjects were coached in the use of both the HUD, for control and monitoring functions, and the TV equipment until they were familiar with their tasks. The subjects then undertook training runs in which they experienced all of the various working conditions.

The experimental design required each subject to complete 16 runs under each of the four working conditions. Preliminary work had indicated that four runs in succession did not cause the pilot any fatigue which could have influenced his performance. Therefore, the subjects conducted four runs at a time, which occupied about 30 minutes, followed by a 30 minute rest outside of the cock . Subjects completed questionnaires during and at the end of the trial when they were also debriefed by the experimenter in order to amplify the questionnaire responses.

The principal dependent variables that were recorded during the experiment can be listed as follows:

TV Tasks

TV target detection time:	Time from start of target bright-up to pi... pressing button on control column.
TV tracking error:	Integral of error between target and cross-hairs in display x and y for 10 second epoch.

HUD Manual Control

HUD tracking error:	Integral of error in azimuth and elevation in display x and y for first and second halves of runs.

HUD Monitoring Task

HUD signal response time:	Time from on-set of HUD signal to subject pressing button on control column.

Eye Movement Data

Time spent looking at HUD:	Pen record of vertical eye movements based on electro-oculographic recording analysed to show percentage of time looking at HUD.

Subjective Data

Opinions on trial and working conditions:	Based on questionnaires and final debriefing.

RESULTS

Only the data from the eight pilot subjects will be considered in this section; the data from the four non-pilot subjects will be considered in the 'Discussion' section.

Effects on TV Tasks of Concurrent Workload

The TV target detection times showed that the addition of the HUD monitoring task increased the average detection time by 0.3 seconds whereas the addition of the HUD manual control task increased the detection time by 0.9 seconds compared to the baseline figures obtained when the subject was performing the TV tasks only. Analysis of variance showed that these differences between conditions were statistically significant at the 0.1% level.

The mean TV display tracking errors over the 10 second period (in arbitrary units) are shown below:

Working Condition	Mean Tracking Error Score	
	in x	in y
TV + HUD Manual	39.0	38.9
TV + HUD Monitor	38.0	23.1
TV only	36.1	22.5

Analysis of variance showed that the differences between the working conditions for the y axis were significant at the 0.1% level.

Effect on HUD Tasks of Concurrent Workload

The HUD manual tracking errors in x and y were compared when the pilot either had or had not the TV tasks, that is the second halves of the HUD manual and the TV + HUD manual runs were compared. The mean tracking error scores are shown below (in arbitrary units):

Working Condition	Tracking Error Score	
	in x	in y
HUD Manual	49.1	21.2
TV + HUD Manual	95.7	55.6

Analysis of variance showed that the addition of the TV task significantly impaired HUD tracking performance in both x and y ($p < 0.1\%$ for both axes).

The response times to the HUD monitoring signal were examined to see if there was any change in mean response time when the TV tasks were required of the pilot. The mean times for the 256 signals are shown below:

Working Condition	Mean Reaction Time (secs)
TV off	0.43
TV on	0.93

Analysis of variance showed that this difference was significant at the 0.1% level.

Eye Movement Data

The percentage of time that the pilot spent looking up at the HUD was analysed according to whether the television monitor was on or off for each of the two working conditions; the results are shown below:

Working Condition	Percentage of time looking at HUD	
	TV off	TV on
HUD Manual Control	60.3	29.3
HUD Monitoring	43.0	21.4

Analysis of variance showed that the differences associated with the TV monitor being on or off were significant at the 0.1% level. In addition, the differences between the two HUD working conditions were significant at the 1% level.

Subjective Data

The questionnaires asked about the conduct of the trial, in particular as to whether the subjects thought that the amount of familiarisation and training that they had received was adequate and whether the work/rest scheduling of the trial had caused any fatigue or eyestrain. Responses to these questions indicated that the subjects thought the training scheme had been adequate (a finding supported by the objective data analysis) and that they thought that their performance had not been affected in any way by fatigue or eyestrain.

Subjects were asked to rank the working conditions from hardest to easiest; the overall order that was obtained from these rankings was (hardest working condition first):

1. TV tasks plus HUD manual control.

2. TV tasks plus HUD monitoring.

3. HUD manual control only.

4. TV Tasks only.

A coefficient of concordance showed that there was significant agreement amongst the subjects on this order of difficulty ($p < 1\%$).

DISCUSSION

Dual Task Performance

When the two tasks of controlling the aircraft and of operating the TV guided missile were performed at the same time, performance on both tasks was degraded compared to the condition when only one task was performed. This general deterioration in performance took place in spite of the instructions to the subjects which emphasised the primary nature of the aircraft control tasks and would also appear to be against the experience of the subjects who, as experienced pilots, would be aware of the relative importance to be given to the two tasks in the real airborne situation. This finding fits in with the work summarised by Rolfe where in the secondary task experimental situation it is customary to find a degradation of performance at the primary task even when the instructions stress the importance of maintaining the highest level of performance at that task.

The subjective ranking of difficulty of working condition showed that the situations in which the pilot was required to undertake both the TV tasks and the HUD tasks concurrently were judged to be the hardest. The order that the subjects unanimously agreed on was that the addition of the HUD manual control task made more difference than the addition of the HUD monitoring task. This finding is confirmed by the effects that the two HUD tasks had on the concurrent TV tasks. For both TV target detection and TV target tracking, performance was more degraded by the need to undertake the HUD manual control task than by the need to perform the HUD monitoring task.

The eye movement data show that when the TV display was on, the subject spent significantly less time looking at the HUD. It is interesting to note that when the TV display is on and the pilot looks at the HUD for only half of the time he spends when the TV display is off, that his HUD tracking errors and the HUD monitoring response times double in size. It would appear from the data on viewing times and on tracking performance in the HUD manual control runs that subjects achieve their best level of performance when they spend some two-thirds of their time looking at the HUD. For the tasks studied in this simulation situation it appears that any other cockpit task that occupies more than the remaining one-third of their time would appear to cause a deterioration in their primary control task performance.

Comparison of pilot and non-pilot subjects

The results obtained from the four non-pilot subjects were very similar in the pattern of performance to the results obtained from the eight pilot subjects. The non-pilot subjects showed the same significant increase in TV target detection times and in TV tracking errors as the concurrent workload was increased. Similarly, performance at the HUD tasks was degraded when the subject was required to undertake the TV tasks.

The main difference between the two groups was that the non-pilots showed a greater fall in performance at the aircraft control task as the total workload was increased. This can be illustrated by the HUD manual control performance scores:

Working Condition	Mean integrated error score (arbitrary units)			
	x axis		y axis	
	Pilots	Non-Pilots	Pilots	Non-Pilots
HUD Manual only	49.1	61.4	21.2	23.2
TV + HUD Manual	95.7	153.3	55.6	64.1

This result agrees with the finding that a secondary task has a greater effect on primary task performance for the less-skilled subject than for those who have achieved a high level of skill on the primary task (see Baker et al, 1951).

CONCLUSION

The increase in workload resulting from the addition of a secondary task in a simulated flying situation resulted in a degradation in performance at both of the tasks despite the instructions given to the subjects.

The subjects did not attempt to maintain their primary task performance level constant and fit in the secondary task when possible but rather tried to achieve a satisfactory level of performance at both tasks. It can be argued that in the real flight situation the pilots would employ different tactics to those they used in the flight simulator. However, no evidence, either objective or subjective, is available from this study to support this hypothesis.

The overall conclusion is that if the highest level of performance is required of a pilot then he should be freed from the requirement to undertake any possible conflicting concurrent tasks.

REFERENCES

1. BAKER, K.E., WYLIE, R.C. & GAGNE, R.M. 'The effects of an interfering task on the learning of a complex motor skill'. J. Exp. Psychol, 41, 1-9, 1951.

2. ROLFE, J.M. 'The secondary task as a measure of mental load'. Chapter 2:9 in Singleton, W.T., Fox, J.G. & Whitfield, D. "Measurement of Man at Work". London, Taylor and Francis. 1971.

DISCUSSION

GUIGNARD When the subject was monitoring the head-up display alone was he spending only 43% of his time looking at it? Is this not rather low? Have you any explanation?

CORKINDALE We did not record in detail where the subject looked in the cockpit. We recorded only the amount of time they spent looking at the head-up display. The rest of the time they looked at the television display or at the other controls.

JEX Was the pilot using only the head-up display to control the aircraft or were other instruments present which he could look at?

CORKINDALE The subjects used other instruments as well.

QUESTION The instructions did not say that making a hit was ... object of the mission. It would seem to me that under these conditions they would try to pay a lot of attention to the television missile control test as long as the aircraft was not getting out of control.

CORKINDALE That's right. Obviously you come back eventually to the difficulty that any simulator study which purports to be a high speed low level situation lacks threat to the subject.

QUESTION I would argue that in fact it was optimal under the ground rules you impose on the subjects and the conditions you impose on them.

CORKINDALE It was optimal under the ground rules the subjects imposed on themselves which was their rewrite of our instructions. The subjects always appear to try to optimise everything or optimise the complete mix. One would rather get pass marks on everything rather than high marks on one task and zero marks on another.

WEWERINKE Did you instruct the pilots to perform the monitoring task?

CORKINDALE We did not formally instruct them on the scan pattern for monitoring the head-up display.

WILKINSON I wonder if you would agree that there may be some high load situations where two tasks may be better than one?

CORKINDALE If you have two tasks, one of which is undemanding or unarousing, then an additional task is likely to improve performance. The difficulty comes when you have two tasks both of which require a fair amount of attention. In that case you may overload the individual.

MIDDLETON You were providing the subjects with two separate tasks practically identical. If you want to get the missile on the target you require a different type of display from the head-up one.

CORKINDALE This has been suggested. If you could in this sort of situation bring about a true integration of the two information sources then this may be one way of getting the pilot to integrate the two tasks. Unfortunately there is doubt as to how far such integration is feasible since the subjects would tend in many cases to regard a so called integrated task as being composed of two elements.

BRUMAGHIM We have been looking to-day at the use of simulation and the emphasis so far has been quite different from my experience. Simulation has been a very effective tool in the research and development of controls and displays.
In this approach there is no need to duplicate all aspects of reality. We have been looking at simulation from one of its weakest points ie the need to duplicate reality. But you can work in a cost effective manner with simulation which will lead to flight validations.

MacNAMARA Thank you for your comment, personally I would agree. One of the problems I have felt with regard to simulation is transferring in an absolute sense as opposed to a screening approach.

SIMULATION OF HIGH WORKLOAD OPERATIONS IN AIR TO AIR COMBAT

Dr. F. M. Holden, Major D. B. Rogers
and Dr. C. R. Replogle

Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433

INTRODUCTION/SUMMARY

General methods and techniques for predicting the mission effectiveness of candidate systems in the preliminary design stage and predicting the human operator's subjective preference for the candidate designs are not generally available. For those techniques which have been used, with some success, there exists the question of validation and general applicability. Thus, tools and techniques which can measure human performance and its consequence upon mission effectiveness are currently limited in their application to specific systems. This report discusses the problems associated with workload measurement, provides a structure for the study and analysis of human performance and describes the tools and techniques used by the 6570th AMRL/EM to provide effectiveness versus design data with specific examples from air to air combat man-in-the-loop simulations performed in the Laboratory. The report concludes with a discussion of man-in-the-loop simulation as a technique for system specific human performance data and as a source for the data required to develop general methods and techniques for predicting the performance of manned weapon systems.

WORKLOAD

Attempts to measure human workload suggest that workload per se is a useful concept in the design and evaluation of system performance. Workload has thus become a figure-of-merit commonly used to rate the excellence of a system or to comparatively evaluate the excellence of several systems. Current measures of workload, however, are primarily dependant upon man's capability to perform, e.g. response/reaction time, his subjective evaluation of the job and its difficulty, e.g. pilot rating schemes, and physiologic measures of effort, e.g. heart rate, oxygen consumption. There are very few measures of workload which arise solely from an engineering analysis of the system itself or from direct measurement of system performance. Thus, workload, while being an intuitively desireable and often used criterion for evaluation and comparison collapses into an situationally dependent, often inconsistent and essentially indirect measure of man's effectiveness and his subjective evaluation of that effectiveness.

SYSTEMS ANALYSIS

We have developed an alternative approach to systems studies involving manned weapon operations. In order to provide a framework for the discussion of high workload operations in air to air combat, the systems using air combat missions will be disected into three parts and an analysis of workload factors for each of the three parts will be done from the viewpoint of the aircraft system designer.

There are three areas of concern to the system designer in the large sense. First is the engineering performance of the system hardware. In air to air combat these factors include climb/dive rate, rate of turn, thrust to weigh ratio, payload, and mission range. Second is the man's ability to operate and control the system. Again in air to air combat these factors include control and display interfaces, control system design, cockpit configuration, life support system, ride and flight control quality. The third major area is that of mission performance. The major factors here in air to air combat include air combat maneuverability, survivability of the aircraft, and the effectiveness of the aircraft in delivering weapons, avoiding threats, en route navigation, take-off and landing. Levels of workload are created or established in all three areas. The engineer designs the hardware system which man must use thus building into the system hardware dependent workload factors. The pilot touches, manipulates and controls the various interfaces available to him according to his concept of system operation and mission objective thus generating on man dependent workload factors. The pilot and his aircraft work together to achieve a mission driven goal and objective thus establishing mission dependent workload factors.

Obviously, all three of these areas interact and are mutually dependent upon each other. Further, it is usually the mission goal which drives the design considerations of the system and also drives the motivation of the human operator to best utilize his expertise and the capability of the system he is operating to accomplish the mission goal. There are, in addition to the above factors, constraints imposed upon the system designer and human operator which include costs, maintainability, reliability, training, and prior experience.

DESIGN CONSIDERATIONS

This representation of the air combat scenario provides a method of putting into perspective some big questions concerning workload. What is workload? How should it be measured? Are there general principles and techniques of workload simulation and measurement which can be applied to any system? Can the application of these general methods and techniques be used in the preliminary design stage to design and build better military systems? Can the uncertainty of human performance and the uncertainty of human response to environmental and psychological stress be included in these general methods and techniques?

The most important short term problem for the system designer is the availability of data to be used in the evaluation of the mission effectiveness of candidate system designs. Valid evaluations of the mission effectiveness of candidate systems can be used in conjunction with costs, maintainability, reliability and training factors to arrive at useful trade-off judgments in the selection of specific design goals.

The most important long term problem for the system designer is solving for the optimal design of a system with tools and techniques which the designer can use himself and in which the designer has confidence. These tools and techniques would include representations of the hardware and human performance and also include representations of the mission scenarios and goals.

The measurement of human workload is obviously involved in the solution of the first, short term, objective, and is obviously involved in the second, long term goal. The short term goal can be satisfied by data alone, while the long term goal requires reduction of the same data into generic predictors of performance. If the techniques used to solve the first short term goal are valid and if in developing and using these techniques the investigators are aiming at the solution of the long term goal then the predictive techniques required in the solution of the second goal will evolve naturally in the process of supplying the system designer with system specific mission effectiveness evaluations. Integral within this philosophy is the requirement for consistency of the techniques and methods used in the evaluation and the requirements that all systems specific studies use consistent techniques and methods which can be validated. These techniques must absorb engineering definitions of hardware systems and essentially similar engineering definitions of human operator performance. Researchers in the area of human workload generally are working towards the second goal since the heart of valid predictors of human and system performance require understanding of and insight into the mechanisms of human and systems operator and interaction. The research appeal of this approach, however, is often ignored and rejected by the system designer who must design, build and test specific hardware to be used by specific types of human operators. Further, the assumptions and constructs required to avoid the vacuum of understanding of mechanisms and data for validation often lead to non-testable results and vagueness intolerable to the hardware oriented designer. Current economic concepts also provide little rationale to support the essentiality of such a long term, reasonably expensive approach.

The structure and philosophy for the simulation and measurement of manned system performance evaluation has thus been logically, although somewhat argumentatively, established. The techniques to be used must evaluate the mission effectiveness of candidate systems. In doing so they implicitly evaluate the workload capability of the human operators. These techniques must also allow the laboratory to collect and structure a data base which can be used to develop general techniques to be ultimately used by the system designer for predicting the performance of candidate manned systems and to optimize the man-machine design interface.

6570 AMRL/EM man-in-the-loop simulations of weapon system engagements is being used by our laboratory to evaluate the effectiveness of countermeasures against the human operator of both ground based antiaircraft artillery and of aircraft engaged in aerial combat.[1-3] Man-in-the-loop simulations are performed in our laboratory to measure the effectiveness of manned aircraft in air combat missions as a function of environmental stresses such as thermal and sustained acceleration loads.[4-5] Finally, man-in-the-loop simulations of air to air combat are performed in our laboratory to measure the mission effectiveness of candidate aircraft cockpit designs.[6-8]

The objectives of the man-in-the-loop simulations performed in our laboratory are to provide the designer with data which he can directly use to evaluate the mission effectiveness of candidate systems or subsystems. Inherent in the effectiveness-vs-design data are the effects of human workload capability since real men are performing the actual jobs required in the real mission. The designer, however, may not be able to discern from the effectiveness-vs-design data whether the differences in effectiveness are due to human workload limitations. The signals recorded during the man-in-the-loop simulations, however, contain all the information required to make this determination.

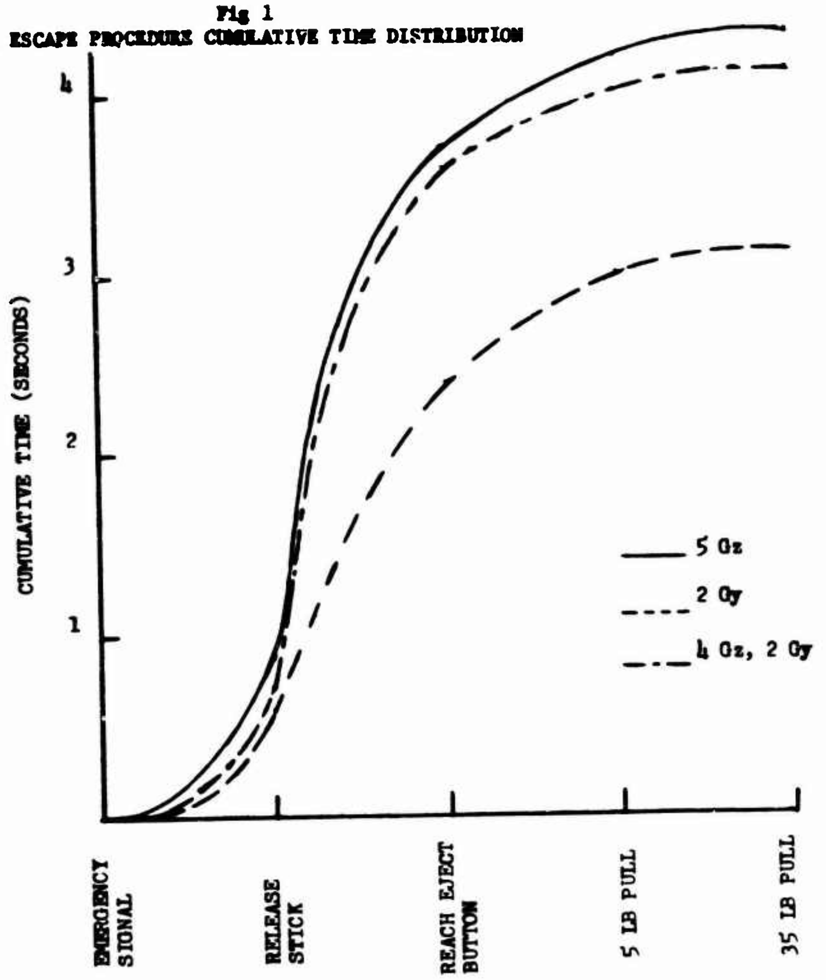
An integral part of the research program in our laboratory is the analysis and modeling of the signals recorded during the man-in-the-loop simulations for the purpose of determining the control capability of man, i.e. his workload capacity and the interactions between man's control capability and the hardware design features and the mission objectives. The analysis and modeling of the human operator is a research effort which will result in generic predictors of human performance. Currently, if the designer wishes to know if the effectiveness of a system in a specific mission is being comprised by human performance limitations, this answer can be provided through the analysis and modeling effort, but only for the specific missions studied in the man-in-the-loop simulations.

In the area of air to air combat our laboratory has provided effectiveness-vs-design data for tilt-back cockpit seat configurations and is currently evaluating two specific aircraft cockpit designs. The laboratory has additionally provided effectiveness-vs-design data for air to air combat countermeasures, and has recently completed man-in-the-loop simulations which provided effectiveness-vs-design data in the area of emergency airborne escape procedures for the B-1 and F-15 aircraft.

B-1

The prediction of pilot performance in flight environments to dangerous to produce in flight test is a generic problem exemplified by a task recently completed for the B-1 System Program Office. The question addressed was that of human performance in an ejection sequence during abnormal flight attitudes. The Dynamic Environment Simulator (DES) was used to produce a man-in-the-loop simulation of the B-1 mission profile. Pilots in a B-1 cockpit simulation flew a low-level, high speed profile, followed by one of a set of adverse flight situations. Acceleration forces as seen in a spin were applied to the pilot who then attempted to perform the required ejection sequence. The study shows (Fig 1) that the time required for the ejection sequence, from start to completion, is determined by the magnitude and direction of the G vector. The results of this study[9] was used by the airframe contractor to evaluate the effectiveness of the emergency aircrew ejection system.

Fig 1
B-1 ESCAPE PROCEDURE CUMULATIVE TIME DISTRIBUTION



F-15

In support of the F-15 High Angle of Attack test program a man-in-the-loop simulation study of an F-15 spin recovery operational sequence was performed. Three F-15 test pilots were subjected to three acceleration profiles. The three acceleration profiles were selected on the basis of the adverse g forces which the F-15 pilots would be expected to encounter in spin recovery operations in the F-15. The profiles lasted for 40 sec and consisted of -4.3 Gx, +1 Gz, -4.3 Gx, -1 Gz and -4.3 Gx with an oscillating Gz component of ± 1 Gz at a .5 Hz frequency. The DES cockpit was configured according to the F-15 cockpit complete with seat restraint system and a mock-up of the spin recovery control box to be used in the actual flight test program. The pilots were F-15 test pilots, and members of the centrifuge subject panel.

The results of this study showed that the pilots could tolerate the three profiles with some incipient nausea reported during the oscillating Gz profile, but that major changes in the F-15 hardware spin recovery system were necessary. The spin recovery control box was modified in its switchology to allow for better switch function and the procedures used in returning the aircraft to a stable configuration in terms of the pilot's use and manipulation of the rudder and control stick were changed.

MAN-IN-THE-LOOP SIMULATION

Man-in-the-loop simulation in conjunction with the evaluation techniques of vulnerability assessment, survivability assessment and weapons effectiveness measurement is a satisfactory technical approach to be used in the evaluation of mission effectiveness of system specific candidate manned systems. Further, data generated during a man-in-the-loop simulation study can be used directly to develop models of human performance and man-machine performance which can be used to optimize the design of systems whose parameters are in the neighborhood of specific system simulated in the man-in-the-loop studies.

The feasibility of using man-in-the-loop simulations for mission effectiveness measurement is dependent upon the resources available to the designer or laboratory and upon the validation of the effectiveness measurements. The advantages of man-in-the-loop simulation should include; a realistic work environment, i.e. real man or man performing the real jobs; a task and task sequence which can be controlled by the experimenter; stressors associated with the real world which can be applied dynamically; a sufficient statistical data base. Also, the human operator should be given weapon performance feedback information and the human operator should be allowed to optimize his control and strategy. The costs of obtaining the mission effectiveness data weighed by the advantage of having a sufficient statistical data base are less than comparable field and flight test data and finally, the real world data base is directly comparable to the man-in-the-loop data base thus validation is possible.

Despite this impressive list of the advantages there are real problems associated with using man-in-the-loop simulation. Funds and manpower resources may limit the detail used in the simulation, and it is not always possible to retionally priortize simulation requirements. Thus an investigator may choose to compromise the fidelity of a display in favor of a more faithful representation of the control system.

These compromises may and usually do lead to some uncertainty in the interpretation of the results of the simulation studies. Engineering equations describing the dynamic response of a target or an airframe or the associated vulnerability representative of the target and the ballistic characteristics of projectile missiles and warheads may be in doubt or contested among the many experts in these areas. Thus one group of researchers may arrive at results which are in conflict with another researcher's results solely because the two groups used differing missile models. The gathering of a statistically significant data base from man-in-the-loop simulation requires good expertise from the discipline of psychology, especially in the areas of experimental design and analysis, subject selection and training, and motivation. Mission effectiveness evaluation is straight forward when the primary concern is weapon delivery accuracy. The significant variables such as hits on target, probability of kill and bomb dispersion are easy to calculate and analyze. If however, the primary concerns are in the areas of maneuvering performance, vigilance, and optimal force mix, engineering definitions of candidate metrics are much more difficult to construct, and availability of differing viewpoints leads to unresolvable compromises and endless discussion.

The problems associated with man-in-the-loop simulation can usually be resolved through discussion among the representatives of the several disciplines required to complete an acceptable man-in-the-loop study. A major problem area in man-in-the-loop studies which receives much discussion but little activity is that of validation. Consider the spectrum of simulations starting with the large number of available digital simulations for everything from antiaircraft artillery/aircraft engagements to full scale campaign models, and the fewer but still impressive man-in-the-loop simulations, to the sparse but extremely expensive field and flight test ranges all of which produce comparable data bases and all of which attempt to replicate environments of concern to the system designer. There are very few examples where a program office responsible for developing a system fully uses the power and resources of these three simulation facilities. Man-in-the-loop simulation data can and should be used to develop better digital models of human and system performance and data from field and flight test programs should be used to validate the man-in-the-loop simulations. The marrying of data sources and objectives of these three areas of simulation will lead to the development of valid techniques for predicting human and system performance.

REFERENCES

1. Holden, F. M., R. Carr, C. R. Replogle, and R. Watkins, "USAF/USA Cooperative OCM study; Final Report (U)", 1971, AMRL-TR-71-16, Wright-Patterson AFB, Ohio.

2. Replogle, C. R., F. M. Holden, R. E. Van Patten and R. Carr, "(U) Compass Ghost flight test simulation and results", 1971, Proc. of 17th Annual Electronic Warfare Planning Conference, Fort Bliss, Texas.

3. Rogers, D. B., C. R. Replogle, F. M. Holden, G. Potor, C. N. Day and G. C. Mohr. "Techniques for dynamic selective simulation under subject control of air to air combat maneuvering", 1972, presented at the Aerospace Medical Association Meeting, Bal Harbour, Florida.

4. Replogle, C. R., F. M. Holden, R. E. Gold, L. L. Kulak, F. Jonas, and G. Potor, "Human operator performance in hypoxic stress", 1970, AMRL-TR-70-122, Wright-Patterson AFB, Ohio.

5. Holden, F. M., D. B. Rogers and C. N. Day, "Models for PA_{O2} and heart rate in humans exposed to acceleration stress", 1973, AFOSR Conference.

6. Rogers, D. B., A. B. Ashare, J. W. Frazier, V. D. Skowronski, K. A. Smiles and F. M. Holden, "Acceleration stress effects on weapon system performance using modified seat back angles", 1973, presented at the Aerospace Medical Association Meeting, Las Vegas, Nevada.

7. Rogers, D. B., F. M. Holden, C. R. Replogle, G. Potor, G. C. Mohr, and C. N. Day, "Performance measurement using pilot controlled Gz maneuvering with a simulated operational task", 1972, given at AGARD/NATO Conference, Brussels, Belgium.

8. Replogle, C. R., C. N. Day, F. M. Holden and D. B. Rogers, "Practical problems in using human operator performance data", 1972, Crew System Design Conference, Los Angeles, California.

9. Adams, W. J., "B-1 ejection controls centrifuge test data analysis", North American Aviation, NA-73-706, TP/PS No. XP 0667-2-001.

PILOT LANDING PERFORMANCE UNDER HIGH WORKLOAD CONDITIONS

Dr. C. A. Brictson
Dunlap and Associates, Inc.
7730 Herschel
Suite AA
La Jolla, California 92037

SUMMARY

A longitudinal study of pilot carrier landing performance was conducted to describe the influence of prolonged operations on pilot performance. A landing performance criterion previously validated in a fleet environment was used to measure and compare pilot and squadron performance variations over time. Three levels of cumulative workload were defined to evaluate concomitant changes in performance associated with each workload. Pilot landing performance improved over time with more improvement found in night performance than day. The influence of practice on carrier landings is discussed in relation to high cumulative workload. The performance criterion was used to identify potential night pilots on the basis of landing proficiency. High and low proficiency pilots also were identified and diagnostic training information provided. A statistically significant increase in night landing performance during high cumulative workload may be due to practice effects as well as workload.

INTRODUCTION

The effects on pilot performance of prolonged high-intensity flight operations and the accompanying accumulation of stress and fatigue are of prime interest to operational commanders. In the U.S. Navy the effects of extended flight operations on pilot performance would be useful for flight frequency and ship deployment schedules as well as for indications of pilot performance effectiveness over time.

Landing performance data seem well qualified to serve as one indicator of pilot performance effectiveness in fleet operations. Carrier landing is acknowledged as a critical phase of flight and is regarded as one of the most complex and demanding tasks required of the naval aviator. The assumption made in this report is that pilot performance in a carrier landing task can be measured and used to describe the relative effects of different intervening variables including workload. The recent development of a valid and reliable criterion of pilot carrier landing performance has facilitated the collection of performance data in the fleet environment and allowed us to measure and describe some of the implications for carrier aviation.

Early studies of pilot landing performance in a fleet environment have focused on relatively short term flight operations usually of one to two weeks duration (1, 2). Ample pre-operation rest periods were available to pilots prior to those studies and any intensive look at performance variations due to cumulative high workloads was precluded by the limited data collection time span.

To complete a longitudinal study of pilot performance at least two things are required. First a criterion measure is necessary that will reliably reflect variations in pilot performance across prolonged time periods of flight operation and different levels of flight workload. Second, different cumulative workload levels need to be operationally defined in order to classify and measure specific segments of work activity in the fleet setting.

A pilot performance criterion that is capable of providing descriptive information on relative levels of pilot landing proficiency has been developed and tested under short duration flight periods. The criterion is called the Landing Performance Score (LPS) and has been found to be a sensitive and valid cri-t[illegible]n of pilot carrier landing performance (3). The LPS has been further validated in two studies d[illegible]ned to predict pilot performance. One study used academic and flight training measures to predict i[illegible] pilot carrier landing qualifications (4). Another study successfully predicted landing performance f[illegible] measures recorded during final approach to night recovery (5).

Meanwhile, in the area of aircrew workload, Nicholson (6) and McHugh (7) have provided the basis for and operationally defined pilot workload level. Nicholson suggested the idea of continuous operational performance capability as distinct from optimum performance demands experienced under high workload and adverse circumstances. McHugh reformulated the idea into a cumulative workload concept derived from flight hours/pilot/day over time during dangerous missions over hostile terrain. He defined three levels of cumulative workload and found them useful in interpreting the results of a study of biochemical and emotional effects of high workload. The successful application of LPS in short duration descriptive and predictive studies and the concept of cumulative pilot workload levels provided the techniques required to describe, analyze and compare the longitudinal pilot landing data presented in this report.

METHOD

Sample

Two squadrons of naval aviators flying the F4J fighter aircraft were used as subjects in the field study. The sample consisted of twenty-nine pilots who embarked on an attack aircraft carrier for a long-term deployment cruise to the Western Pacific. All pilots were carrier qualified and ranged in rank from Ensign to Commander with a mean age of twenty-nine years. On the average pilots had 1.25 years of carrier experience, 1451 hours of general flying experience and 706 hours of specific flight time in the

F4 aircraft. The two squadrons were considered representative of the existing population of naval jet pilots and in terms of mean age, education and flight experience thought to be an unbiased cross-sectional segment of the naval fighter pilot community. Two A7 squadrons also were used in the study to provide comparative data for a different aircraft type. The A7 is an attack aircraft with a somewhat different mission than the F4 and was selected as a second sample for additional information on performance changes over time in different aircraft types. The A7 sample consisted of 22 pilots who completed the entire cruise. Their background and experience were similar to their F4 counterparts.

Data Collection

Landing performance data were collected for different time periods during the deployment. In some cases data were also recorded during predeployment refresher training to obtain additional baseline performance information. Landing data for day and night approaches were obtained from ship logs and Landing Signal Officer (LSO) log books and included information on pilot, aircraft, launch and recovery time, landing sequence, landing results and LSO qualitative evaluations of each recovery. Environmental data in the form of wind-over-deck, sea state, weather, deck pitch and trim, ceiling and visibility and sun/moon/horizon data were also routinely recorded.

Landing data were categorized by individual pilots, squadron or aircraft type, and by day or night recovery. Each aircraft approach resulted in one of six landing categories: a technique wave off, wire #1, 2, 3, or 4, or a bolter (touchdown without arrestment). The landing categories were assigned weighted scores representing the quality of landing as described by Brictson (3). In simple form each landing category was objectively assigned a score based on a scale of 1 to 6. The scale had been previously rank-ordered and assigned weights by senior LSOs (see Figure 1). The result was a criterion measure of pilot landing performance called the Landing Performance Score (LPS). The LPS is based on an equal interval scale of landing quality that represents an LSO concensus of the relative numerical value of each possible landing outcome.

Each pilot carrier landing was transformed into LPS data and converted to card form for computer statistical treatment and printout. LPS data were compiled into specific time periods and statistically grouped to facilitate comparisons ove relatively equivalent time segments.

Pilot Workload

Workload considerations were based on an operational definition of three levels of aircrew workload as described by McHugh (7). A cumulative work concept was employed based on the product of average flight hours/per pilot/per day, the number of consecutive days of flying activity and the relative danger of the mission. This conceptual framework resulted in the definition of three cumulative workload periods: zero, moderate and high. The zero cumulative workload level consisted of pilot landings made after a prolonged non-flying period, usually in-port, during which adequate rest, recreation and recovery time were provided. A moderate cumulative workload level was defined by eleven consecutive days of flying missions over nonhostile territory. The high cumulative workload level consisted of double the moderate level plus missions over hostile territory with the possibility of death, capture and/or heavy enemy offensive actions. Thus, high workload was defined by 22 days consecutive flying over hostile terrain with a high degree of danger. The three workload levels were selected in coordination with another Navy data collection team that was investigating biochemical, emotional and sleep correlates associated with stress and fatigue over long term flight operations for the same pilot sample (7).

RESULTS

The results are presented in two parts. Part one describes longitudinal measures of pilot landing performance by comparison of population, squadron and pilot data. Part two describes the effects of cumulative workload on pilot performance.

Longitudinal Measures

Population Performance. Figure 2 illustrates a population distribution of day and night landing performance scores based on approximately eight thousand carrier landing approaches by all fleet jet aircraft. It shows a mean night LPS of 4.55 and a mean day LPS of 4.78 across all aircraft. Comparison of LPS data in this study for all F4 and A7 landings over the entire cruise resulted in the following information.

	Day LPS			Night LPS		
Sample	$\overline{X}$	σ	Percentile	$\overline{X}$	σ	Percentile
F4	5.08	.26	88	4.67	.31	78
A7	5.05	.22	87	4.87	.28	90
Baseline Population	4.87	.24	50	4.55	.25	50

Both F4 and A7 night and day landing performance were above the population mean for all aircraft and signify that a high level of proficiency was demonstrated by the deployed aviators throughout the cruise. In fact the A7 night mean performance was also above the day mean landing performance for all aircraft and fell at the 90th percentile for the night baseline population. F4 night performance was at the 78th percentile for all night carrier landings. Although F4 and A7 LPS means were found to be higher than the population mean neither F4 or A7 differences were found to be statistically significant.

Squadron Performance. F4 and A7 squadron performance for different equated time periods during the cruise is shown in Figures 3 and 4. Figure 3 shows F4 LPS data compared day and night across refresher training and deployment. Day landing performance was consistently above night performance and showed a gradually increasing slope of landing proficiency over time. Although the sixth time segment shows a slight drop in landing performance the gradual upward trend is evident again in the last time period. Overall day landing performance showed a gradual but continuous increase in landing performance scores from refresher training through the end of the deployment.

Night F4 LPS data show a steep early gradient of improvement over time compared to day landings but also reflect more variability across time segments than corresponding day LPS for the same pilots. The finding is consistent with other reported day and night landing differences (1). Time segment five shows a marked increase in performance followed by a gradual decline through time seven. In general night landings improved over time from refresher throughout the cruise.

A7 day and night landing performance are shown in Figure 4. Day landings are characterized by a moderate improvement gradient over time similar to F4 day performance. A7 night data show a steep improvement curve most significantly depicted between time segment one and two but characteristic of the entire cruise. In no instance did average night landing performance exceed day performance for either aircraft type.

To determine whether the gradual pilot landing improvement across time was a function of increased cumulative landing practice, the first three time periods were analyzed by five day intervals at the start and end of each period. Each five-day period was depicted by a mean LPS day and night. For night landings there was a consistent drop in landing performance between time periods with night scores lower at the start of each time period compared with the end of the preceding time period. Cumulative practice effects for night landings were verified by the data analysis. Day landings appeared to be more consistent with no appreciable change in performance as a function of increased practice or stop and start time.

Pilot Performance. Individual pilot landing performance is described in Figures 5 and 6. Figure 5 shows F4 pilot LPS data rank-ordered by day with each pilot's corresponding night landing performance plotted directly below with his day rank order. Individual day landing performance shows a consistent high level of proficiency within the F4 pilot sample. Night landing data are more variable. Pilots #9 and 10 show the most consistent day/night landing performance while the greatest difference in day and night landing proficiency shows up for pilot #6 and #14. Pilot #14 had the lowest LPS day and night for the entire cruise.

Figure 6 shows individual A7 pilot landing performance rank-ordered by each pilot's day LPS with his night LPS aligned with his day rank order sequence. The most consistent pilot was pilot #1 who had the highest day and night landing performance score for the cruise. Pilots #11, #12, #13 and #14 were most consistent day and night as measured by the smallest difference in day and night LPS values. Four pilots had higher night LPS values than day with pilot #19 showing the largest increase in landing performance at night compared to day. Four pilots had large night landing performance decrements (Pilots #4, #5, #10, #15). Again, as with F4 pilots, day landing performance was relatively consistent across pilots with larger variability in individual night landing performance.

Pilot reliability data indicate that during the cruise pilots were consistent in their landing performance as measured by correlation coefficients. Day and night pilot LPS data correlated for A7 and F4 samples at $r = .53$ ($p < .01$) for the entire cruise. In terms of pilot consistency between time periods three time segments of F4 pilot LPS data were correlated with independently derived LSO evaluations of each pilot's landing performance. The results indicate a statistically significant relationship with correlations of $r = .83$, $.64$ and $.87$ for time segments 1, 2 and 3 respectively. The first and third correlations are significant at the .01 level with .64 being significant at the .05 level.

Figure 7 illustrates the diagnostic capability of the LPS measurement technique for two F4 pilots, #1 and #14, the highest and lowest landing performers for day recovery operations. Wire three is the aiming point in carrier recovery operations and as can be seen in Figure 7 Pilot #1 touches down on the target wire 30 percent more than Pilot #14 during day recovery periods. At night Pilot #1 lands more than three times as often at the target wire than #14 who boltered once in every four night carrier approaches during the cruise. Relative differences in day and night landing outcomes are readily apparent from the histogram. At night Pilot #1 catches the target wire 1/3 less than by day and increases the proportion of his long (#4 wire) and short landings (#1 wire). At night the low performing pilot (#14) shows the greatest increase in long landings (bolters and #4 wire) at the expense of target wire arrestment. Figure 7 represents a diagnostic description of pilot landing performance signatures or baseline performance data for prolonged operational flights.

Cumulative Workload Measures

Different patterns of landing performance by F4 pilots across three levels of cumulative workload are presented in Figure 8. Day and night landing performance scores dropped slightly from a zero workload to a moderate workload which consisted of 11 consecutive days flying over nonhostile terrain. The most significant change in pilot performance appears at night between moderate workload and high workload or 22 days consecutive flying over hostile territory. The increase in night landing performance from moderate to high cumulative workload was found statistically significant at a $t = 10.97$, $p < .01$. Landing performance improved significantly at night under high levels of workload compared with zero and moderate work levels. By day, however, pilot landing performance measured under moderate and high workload levels was consistently below zero workload level landing performance.

F4 night landing capability was highest during high cumulative workload. F4 pilot night LPS data were also found to be higher than either of the performance means for the entire cruise or the baseline data. Day LPS data were highest under zero workload while moderate and high workload landing performance was equivalent to the day F4 mean for the entire cruise and slightly above the baseline norm.

DISCUSSION

Pilot landing performance during long duration flight operations can be measured and described through application of the Landing Performance Score metric. Squadron and individual pilot performance signatures were obtained during fleet operations and variations in performance were described for different time periods including three levels of cumulative workload. The significance of LPS application to actual fleet operations may be in the broad spectrum of operational landing performance that can be quantified, described and diagnostically fed back to all levels of command from individual pilots to fleet commanders. The LPS provides a common, objectively derived performance scale that can be used to develop performance standards for different levels of pilot training and experience or to provide individual landing performance curves for both day and night operations.

There are several operational implications of the study which may have practical significance to the fleet. The possibility that some pilots may land more effectively at night than by day was suggested by some of the A7 pilot performance data. Four A7 pilots had higher night landing scores than day, while 36 percent of the A7 pilots had night scores approximately equal to or above their day landing scores. In addition, it was demonstrated that diagnostic training feedback on the quality of landing performance could be obtained from LPS data. Pilot landing scores were used to rank-order plilots on a continuum of landing quality and to identify high and low proficiency pilots. Pilot landing distributions were compared for diagnostic data such as target wire arrestment or long and short landing trends. Such types of information could be useful to squadron and airwing officers or LSOs to assist in effective utilization, training and assignment of fleet aviators. In view of the current energy crisis or other flight curtailments such data conceivably might be extended to provide operational selection criteria for individual flight and mission assignment, especially if flight frequency shortages affect maintenance of long term pilot landing proficiency.

Different levels of cumulative workload were found to influence carrier landing performance only in the high workload condition for night performance. The increased tempo of operations and accumulated night experience acquired over 22 consecutive days of flying suggests that recent night landing experience may be related to the significant increase in night landing performance during high cumulative workload. Some evidence exists to support the idea that practice effects are interrelated with high cumulative workload, especially at night. During the cruise pilots had a significantly greater number of day landings than night landings so that under high workload conditions the increase in performance may be partially due to more night experience rather than a simple effect of high workload. LPS data for the entire cruise lend added support to the contention of practice effects since landing performance showed a gradual improvement for both day and night performance over time. Night performance tended to show a steeper improvement gradient over time, especially in the early stages of the cruise. Day landing performance changes were more gradual than night changes probably due to a greater number of day landings per pilot for each time period. With an equivalent increase in night landing experience we would hypothesize that pilot night landing scores for zero and moderate workloads would increase to correspond more closely with day results. However, not all of the increase in night performance can be accounted for by practice effects. Some of the striking increase in night performance under high workload is probably attributable to optimum pilot performance as suggested by Nicholson. Our data show a high workload night performance level that is considerably higher than that accountable by practice effects alone.

Night landing performance was isolated from day performance because it is more demanding of the naval aviator. It requires more skill, attention and concentration by the pilot compared to day recovery and probably is more sensitive to fluctuations in cumulative pilot workload. Previous studies (8, 9, 10) have separately reported increased physiological response during periods of pilot final approach and landing. Those data indicate that landing is a demanding and stressful task that can be associated with changes in heart rate, finger tremor and blood chemistry. In this study we found night landing performance to be more sensitive to workload fluctuations than day performance even when practice effects are taken into account. In any future work in this area night landing measures should be used as the primary performance measure to obtain data that is more sensitive to operational workload conditions.

In summary this paper has discussed the application of a criterion of pilot carrier landing performance to describe performance variations associated with prolonged operations and different levels of cumulative workload. The intriguing question of the predictability of pilot performance from intervening variables such as experience and psychophysiological factors collected concurrently with the data reported here is discussed in a sequel to this report (11).

REFERENCES

1. Brictson, C.A. Measures of Pilot Performance: Comparative Analysis of Day and Night Carrier Recoveries. Santa Monica, California: Dunlap and Associates, Inc., June 1966.

2. Brictson, C.A., Ciavarelli, A.P., and Wulfeck, J.W. Operational measures of aircraft carrier landing system performance. Human Factors, June 1969, 11(30) 281-290.

3. Brictson, C.A., Burger, W.J., and Wulfeck, J.W. Validation and Application of a Carrier Landing Performance Score: The LPS. Inglewood, California: Dunlap and Associates, Inc., March 1973.

4. Brictson, C.A., Burger, W.J., and Gallagher, T. Prediction of pilot performance during initial carrier landing qualification. Aerospace Medicine, May 1972, 43(5), 483-487.

5. Brictson, C.A., Burger, W.J., and Kennedy, R.S. Predicting the quality of pilot landing performance during night carrier recovery. Aerospace Medicine, January 1971, 42(1), 16-19.

6. Nicholson, A.N. Aircrew workload. RAF Quarterly, 1972, 12, 273-280.

7. McHugh, W.B., Brictson, C.A. and Naitoh, P. Emotional and biochemical effects of high workload. Paper presented at NATO AGARD ASMP Specialist Meeting, Oslo, Norway, April 22-24, 1974.

8. Nicholson, A.N., Hill, L.E., Borland, R.G., and Ferres, Helen M. Activity of the nervous system during the let down, approach and landing: A study of short duration high workload. Aerospace Medicine, 1970, 41, 436-446.

9. Roman, J., Older, H., and Jones, W.L. Flight research program VII, medical monitoring of naval carrier pilots in combat. Aerospace Medicine, 1967, 38, 133-139.

10. Austin, F.H., Gallagher, T.J., Brictson, C.A., Polis, D., Furry, D.E., and Lewis, C.E., Jr. Aeromedical monitoring of naval aviators during aircraft carrier combat operation. Aerospace Medicine, June 1967, 38, 593-596.

11. Brictson, C.A., McHugh, W.B., and Naitoh, P. Prediction of pilot performance: Biochemical and sleep mood correlates under high workload conditions. Paper presented at NATO-AGARD ASMP Specialist Meeting, Oslo, Norway, April 22-24, 1974.

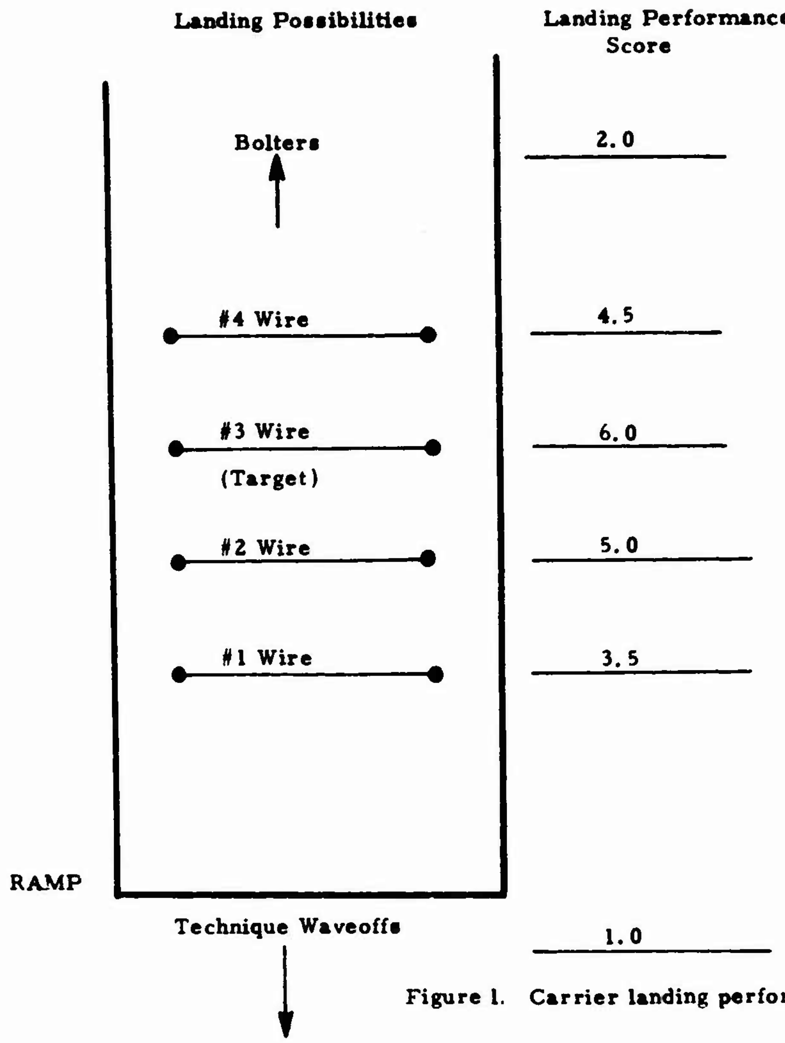


Figure 1. Carrier landing performance criterion

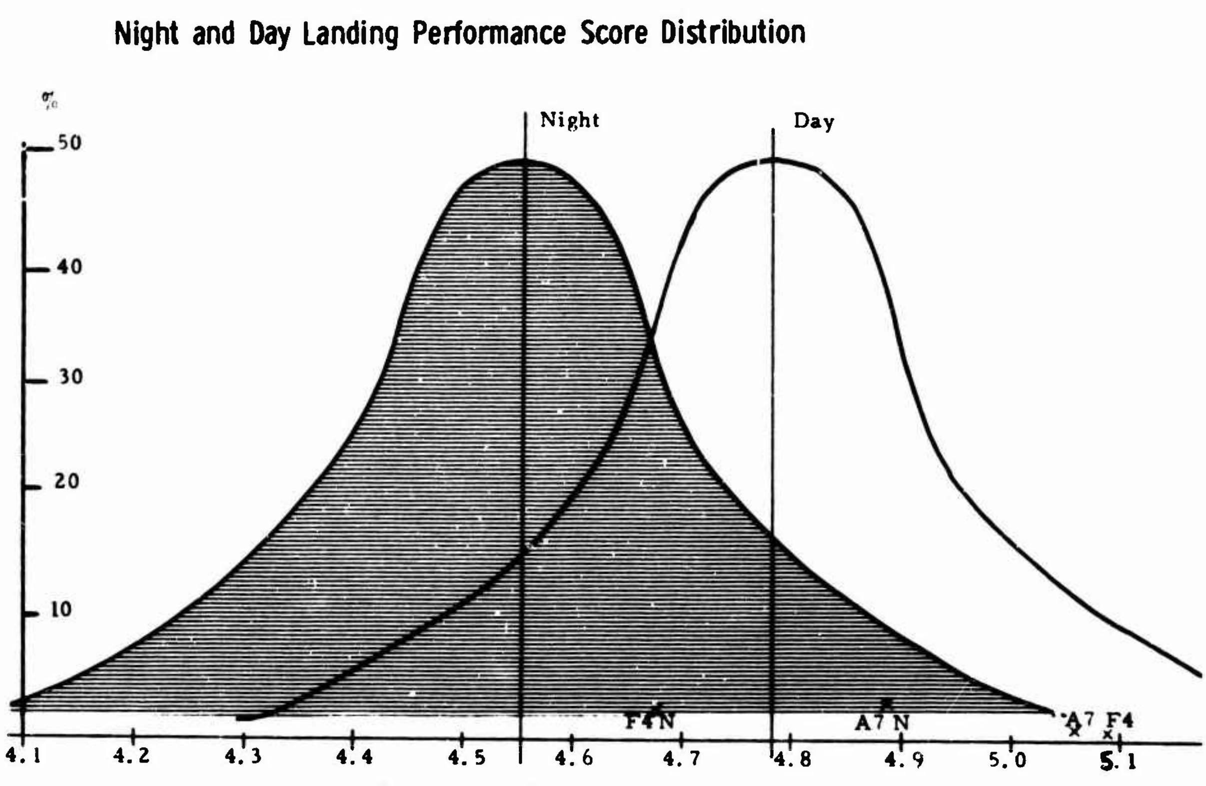


Figure 2. Comparison of F4 and A7 LPS data with population distributions.

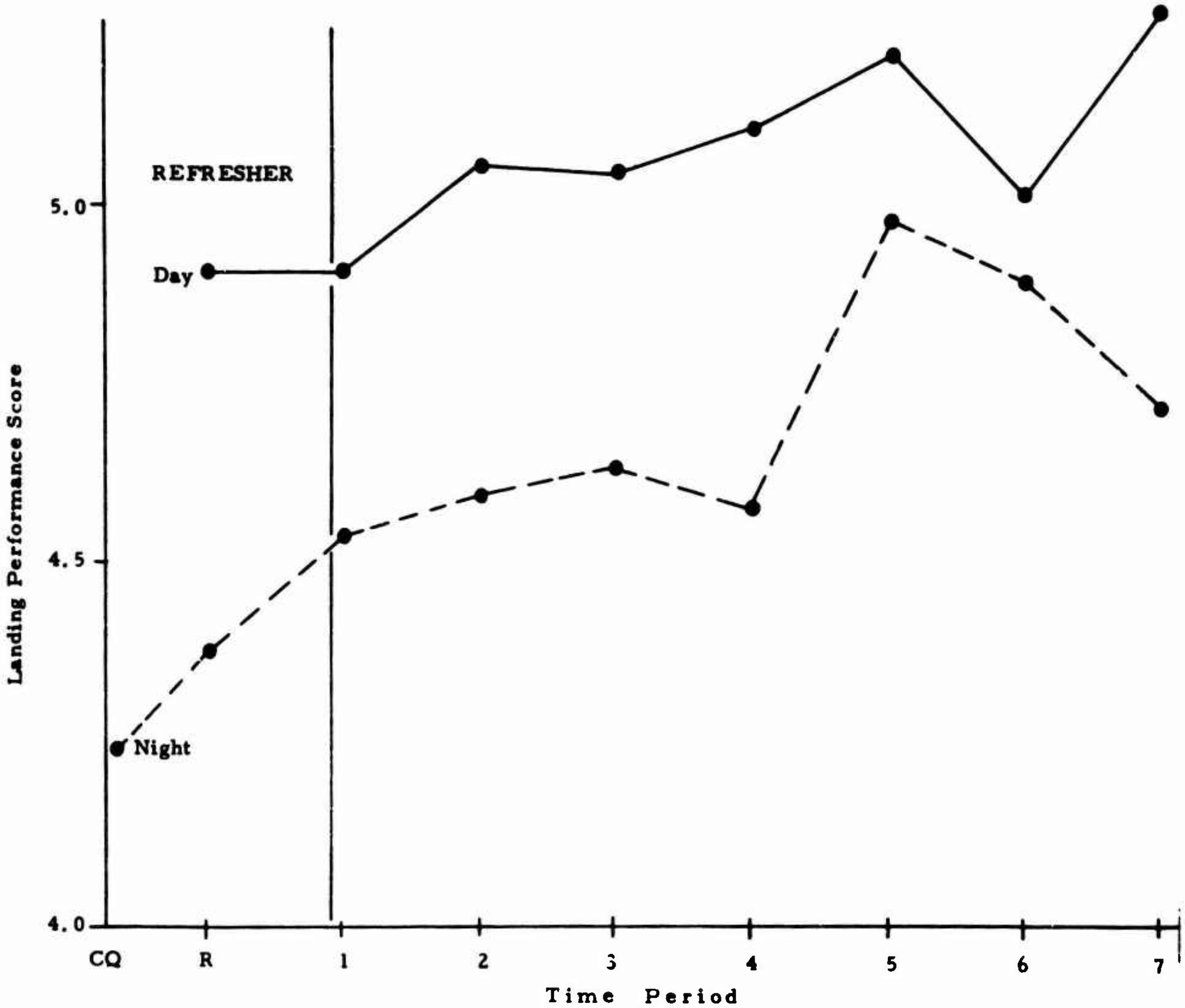


Figure 3. Comparison of F4 day and night landing performance for refresher training and seven time periods.

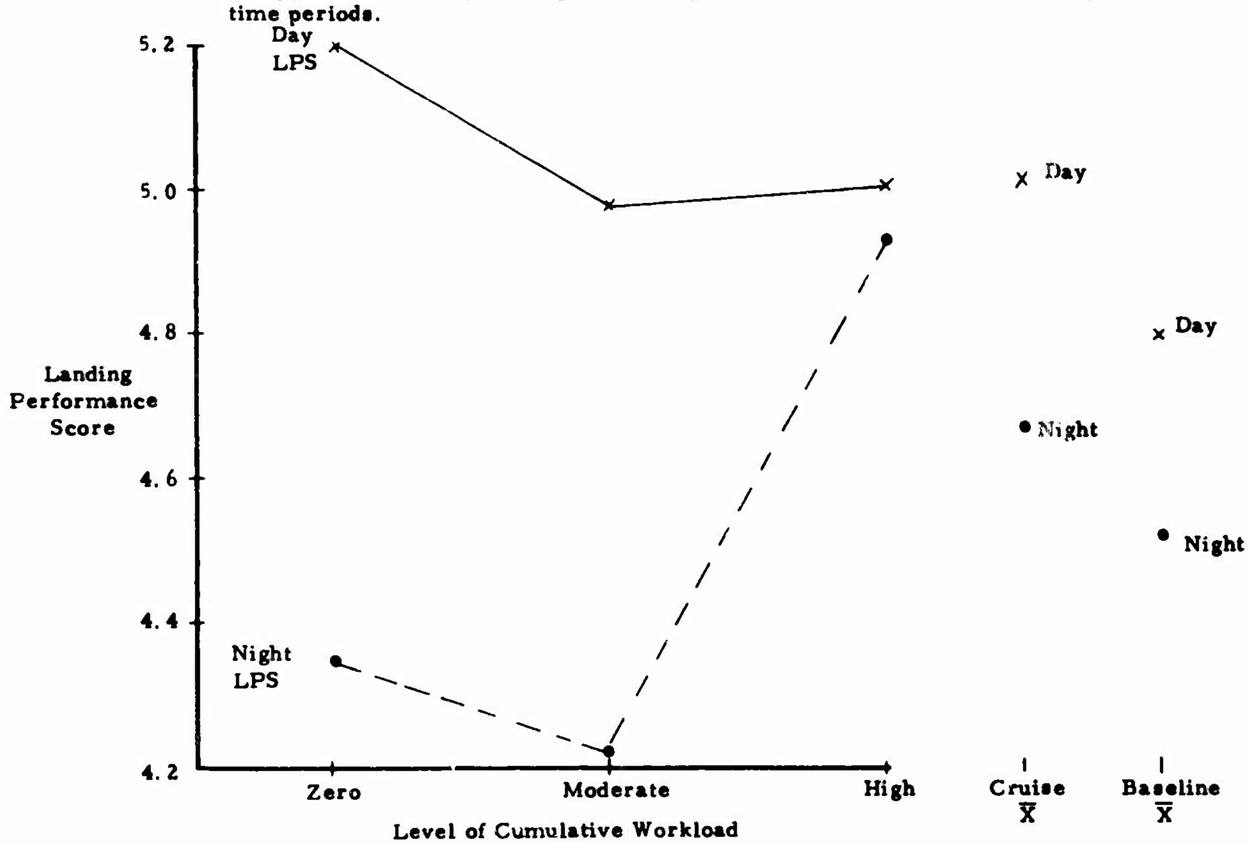


Figure 8. F4 day and night landing performance scores for three levels of cumulative workload.

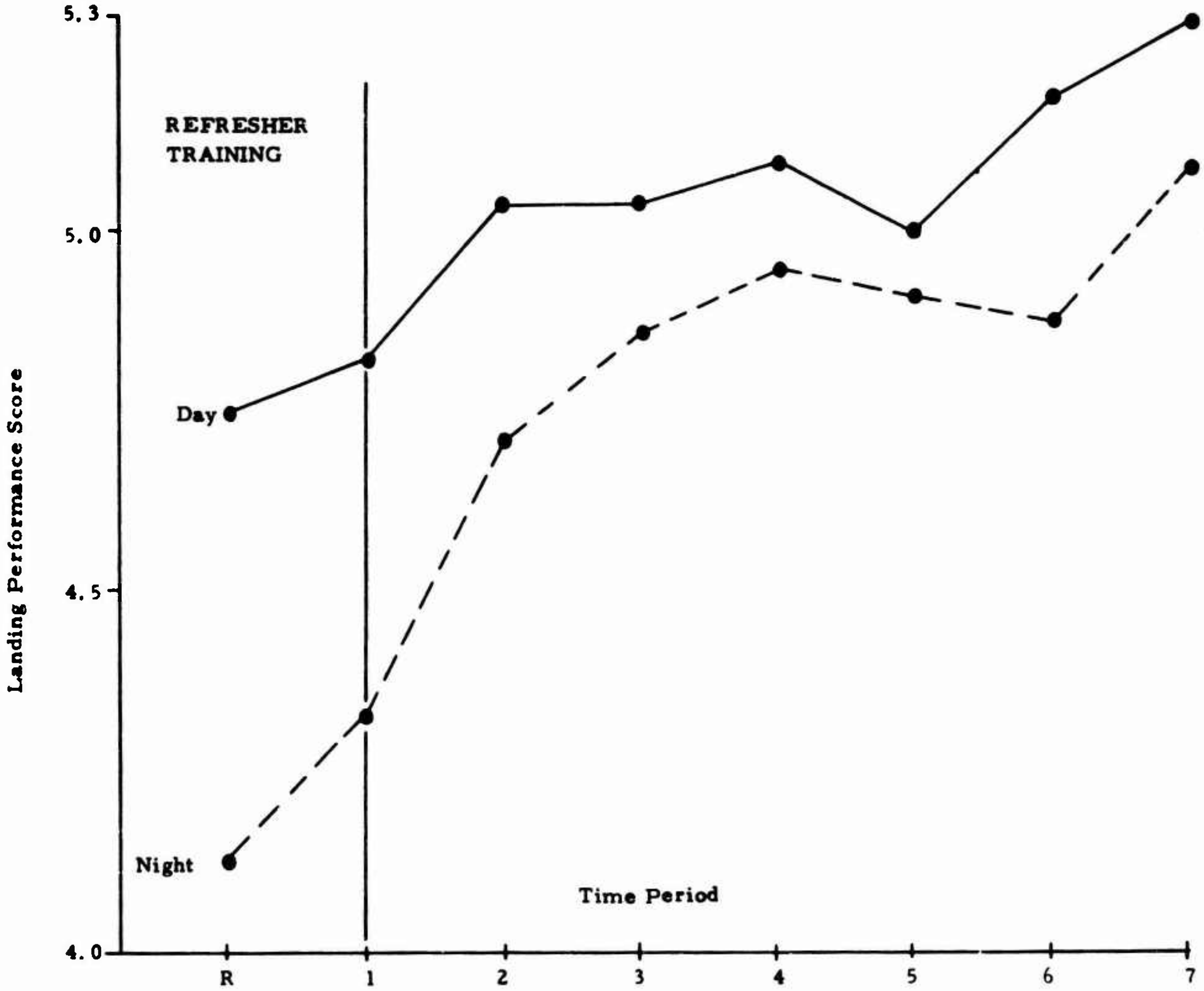


Figure 6. Comparison of A7 mean landing performance across seven time periods.

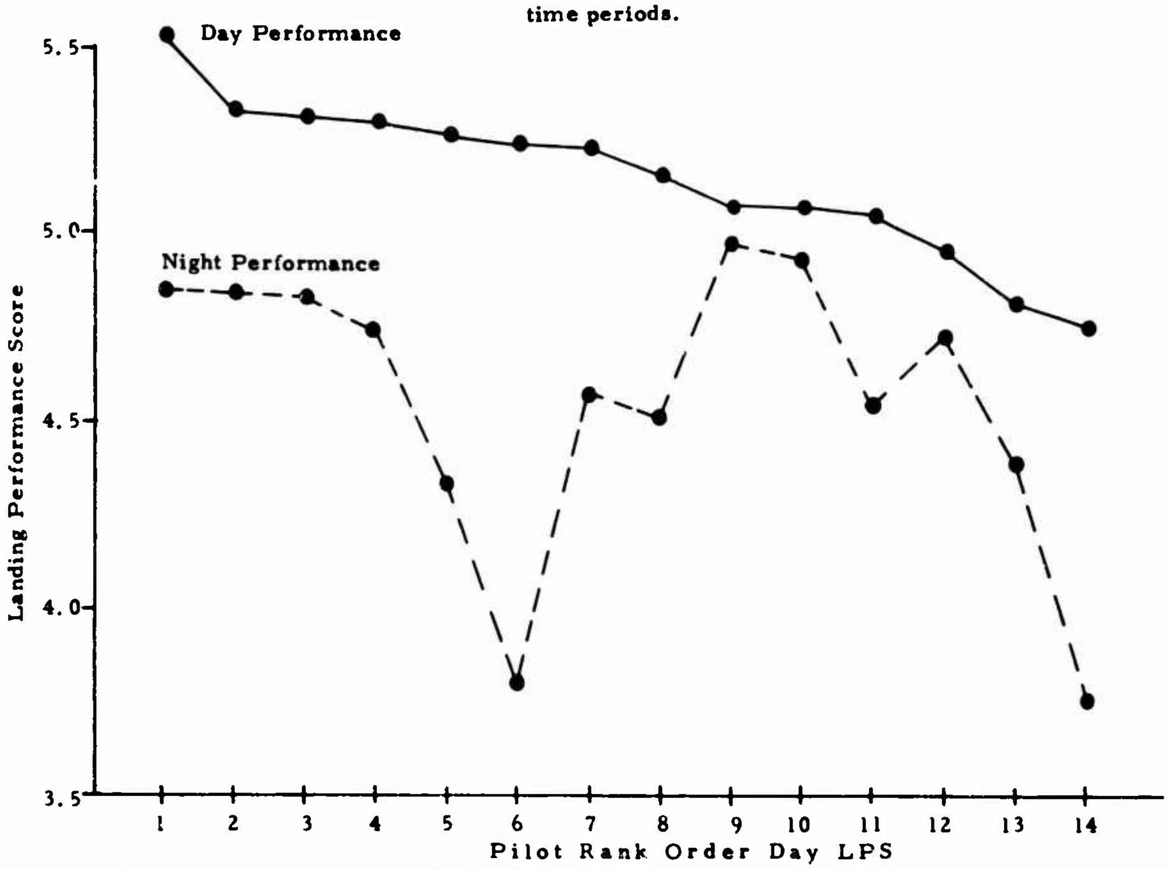


Figure 5. Comparison of F4 pilot X day and night performance scores for entire cruise.

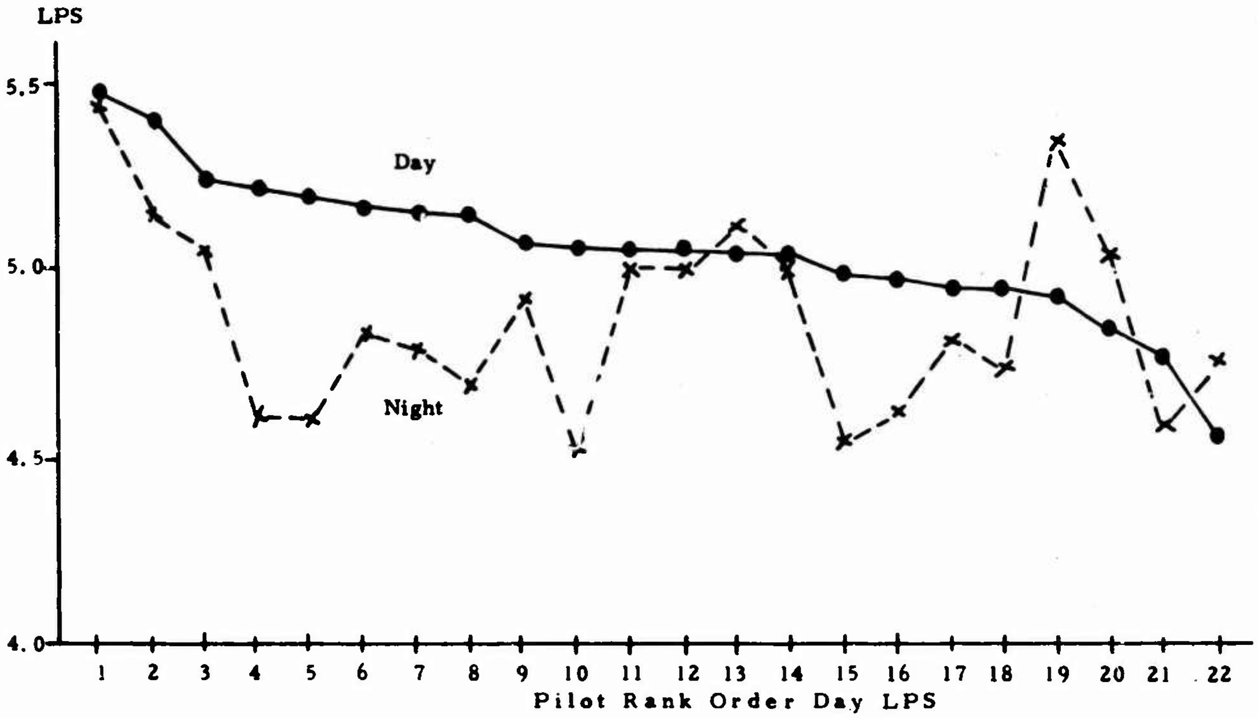


Figure 6. Comparison of mean day and night A7 pilot performance for a nine-month time period. Pilots are rank ordered on day LPS data.

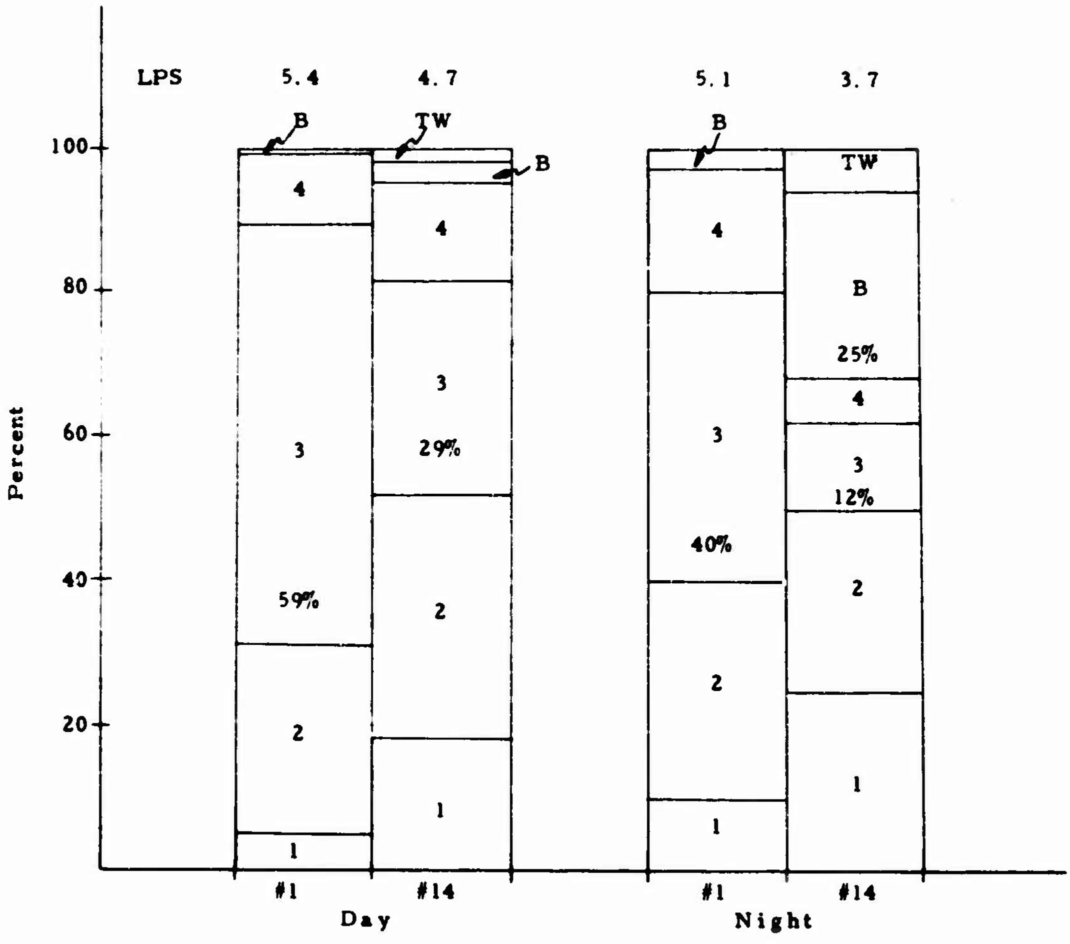


Figure 7. Pilot landing performance signatures. Comparison of high and low pilot LPS data for a nine-month period.

DISCUSSION

JEX	Did you collect information other than on the touch down point. This may help to diagnose poor or good performers from the standpoint of angle of descent, rate etc?
BRICTSON	Not in this particular study. We have collected previously about 6,000 day and night final approach records which describe this type of information. We could account for a large percentage of performance variants associated with the final approach by simply using a final or terminal landing condition. There is a number of reports which correspond this landing performance score to the radar records of the final approach, including glide slope angle, altitude above the deck, sink rate, pitch angle and final approach speed. In addition we had the LSO (landing safety officer) estimates of final approach and landing.
Von GIERKE	What spread in the data do you get through weather conditions?
BRICTSON	During light to heavy rain as opposed to normal recovery conditions we found no significant differences. The greatest difference we find is with a pitching deck.

AIRCREW WORKLOAD AND HUMAN PERFORMANCE
THE PROBLEM FACING THE OPERATIONAL COMMANDER

Lieutenant Colonel W.D. Macnamara
Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue, West
P.O. Box 2000, Downsview, Ontario, Canada
M3M 3B9

SUMMARY

The information available to an operational commander on the reliability and serviceability of the human component in the air weapons system is limited when compared to that available for the aircraft and other systems components. The common use of total flying hours does not provide the commander with information consistent with that now available from aircrew workload and performance studies. A basis for and the background in the development of a trial approach to providing commanders with better information on the human component is described.

INTRODUCTION

For over 60 years, the problems of human performance in aviation have been subject to study in varying intensity. In recent years, the study of aircrew workload and performance has become increasingly important for three main reasons:

a. modern aircraft are more reliable, sophisticated, and efficient and commonly can outfly their crews;

b. airspace and airports have become more congested, increasing the demands placed on crews; and,

c. financial constraints are forcing maximum effectiveness and utilization of both aircraft and human resources.

A relatively recent increased interest in the problems of aircrew workload and performance is reflected in the large number of papers on this subject in the past five years. Notable are those from the Hartman group in the U.S., the Klein group in West Germany and the Nicholson group in U.K. While those of us interested in this area enjoy the understanding arising from the freedom of communication in the literature and at international gatherings, the operational commander has a problem.

The problem facing the operational commander is how to obtain for practical use the human performance data that we, as specialists, have. How can he match anticipated workload with crew performance potential?

It is the aim of this paper to explore the nature of this problem and to present a practical approach that we have started using in the Canadian Armed Forces.

THE NATURE OF THE PROBLEM

The military commander is an executive whose role is to manage his resources to assure the effective attainment of his assigned mission. In air operations, excluding the self-evident moral obligations, the conservation of aircraft and manpower resources is good business. Not only the financial cost but also the time cost in the replacement of the resources could lead to operational limitations that would compromise the probability of successful mission completion. The resources cannot be committed capriciously. The commander must therefore have data on the probability of effective performance of his air weapons systems before he can make his final strategic and tactical decisions leading to the disposition of his resources.

For the aircraft, aeronautical engineering and maintenance technology provide the commander not only with a reliable system but also with a pre-mission indication of the serviceability of that system. Notwithstanding 60 years of effort, we cannot provide him with similar data on the reliability and serviceability of the human component. There is no instrument that can identify for the commander the performance capability of an individual aircrewman, determine whether or not a crewman is suffering from fatigue sufficient to cause a performance decrement, or correlate performance decrement with the probability of successful mission completion. To date we cannot guide the commander, in human performance terms, through that decision pathway in which he must integrate the various priorities of flight safety and mission completion in peace, and of maximum effectiveness at minimum cost, in war

PRESENT POSITION

As evidence that the human performance guidelines for air operations remain somewhat vague, it is interesting to note the following variation in aircrew flying times found in our 1972 survey of seven military air services. (1)

TABLE I
SURVEY OF AIRCREW FLYING TIME REGULATIONS

Air Force	Maximum Monthly Flying Time	Limit on Rate to Monthly Maximum	Maximum Aircrew Duty Day	Minimum Daily Crew Rest
A	120 hrs	50 hrs/7 days	15/17 hrs*	10 hrs
B	125 hrs	Nil	18 hrs	9 hrs
C	120 hrs	Nil	16-20 hrs*	11 hrs
D	120 hrs	Nil	16 hrs	14 hrs
E	125 hrs	Nil	16 hrs	12 hrs
F	90/100 hrs*	25-30 hrs/7 days*	10 hrs	12 hrs
G	125/150 hrs*	50 hrs/wk	12 hrs	15 hrs

*Varies with type of aircraft

The monthly maximum flying times range from 90 hours to 150 hours. The controls on the distribution of the flying hours vary from nil to as little as 25 hours per 7 days. Maximum aircrew duty days range from 10 hours to 18 hours with the minimum daily crew rest from 8 to 15 hours, allowing some adjustments for special circumstances.

Against this rather wide variation in regulations, both past and recent literature recognize the variability in human performance between and within individuals, between and within given days, months and seasons. The masses of data arising from studies of human performance in a wide range of specific aircraft operations permit very few generalizations; but, specific studies could probably be found to support or refute the flying time positions held by any of the above air forces.

In seeking guidelines in how to handle this problem, we conducted, in conjunction with a number of exercises during the period 1971-73, informal interviews of a cross-section of supervisory aircrew in the Canadian Armed Forces to determine their general knowledge of performance, workload and fatigue effects. Out of these interviews only one common thread appeared - the judgement of how well a given aircrewman is performing at a given time is dependent upon how well his normal performance is known by his peers and supervisors. Failing any performance decrement being noticed at the supervisory level, the regulatory control is simply flying hours. The individual still has the right to seek the counsel of the Flight Surgeon if he feels "fatigued", but the responsibility is with the individual. However, several perceptive aircrew encountered in our interviews expressed concern in their ability to recognize any impairment of their personal capability because they also recognized impaired judgement as symptomatic of fatigue.

The recent literature clearly points out that the daily, weekly, or monthly flying hours cannot be used as simple determinants of the probability of next mission success in a given crewmember. Looking at recent operationally oriented studies, Nicholson (2, 3) has emphasized that the quality of sleep is important and that chronic disturbances in the sleep pattern tend to lead to reduced duty capability over a period of days. Hartman (4) has pointed out that although a strenuous mission can be carried out successfully, evidence of the 'physiological cost' is seen days after the completion of the mission. The work of Morgan and Alluisi (5) has indicated that after sleep deprivation, performance may return to near normal after 24 hours rest. Klein (6) has details of some performance decrement during periods in which sleep is normally expected - the so called 'physiological low'. Finally, Harris and O'Hanlon (7) have emphasized the complexity of the environmental, physiological and psychological stresses that may contribute to reduced performance.

From the literature, we concluded that the activities of individual aircrew in the 24 hours immediately prior to their mission could function as reasonable parameters for operational commanders to judge the performance of their aircrew. It was further reasoned that by educating the operational supervisors in the factors that contribute to performance decrement and encouraging them to recognize and manage individual variations in performance, they could manage their men rather than being limited by flying time regulations.

THE APPROACH TO THE SOLUTION

Using a modification of a list of potential stress areas from Harris and O'Hanlon (7) the following approach has been developed for presentation to Flight Safety officers and Squadron Commanders.

Operational Effectiveness = Human Effectiveness

Human Effectiveness = Management of Human Design Limitations

Management of Human Design Limitations =

Research and Development +

Clinical Surveillance +

Education of Operational Commanders

Human Design Limitations are seen in the effects of:

Prolonged Physical Work

Sleep Deprivation

Circadian Desynchronies

Reaction to Infection

Environmental Stresses

- Temperature
- Weather
- Decreased atmospheric pressure
- Noise, vibration, impact, acceleration
- Visual and Vestibular Problems
- Toxic Compounds
- Confinement
- Nutritional, water and electrolyte balance
- Weapons effects (including NBCW)

Situation Stresses

- Social and Domestic activities
- Change in Life Events
- Inexperience
- Threat of Injury or Death
- Command responsibility

Moderate detail is provided to indicate how each of the factors affects human performance in aircrew. Commanders are encouraged to try to identify factors contributing to a perceived performance decrement in a given individual. Emphasis is placed on the need for squadron and flight commanders to know and monitor the normal activities, behaviour, and flying performance of their crews and, particularly, to recognize the individual and daily variations.

The use of this approach has the following objectives:

a. to provide the operational commander with improved information on the performance potential of those under his command;

b. to permit the squadron commanders to manage flying hours;

c. to validate conclusions found in or based on existing literature; and

d. to identify priority areas for study, through feedback from commanders.

Although early indications are that the commanders are interested in this approach, we have as yet no indication of overall aircrew acceptance or co-operation.

CONCLUSION

In presenting this first hesitant step at improving the management system for aircrew workload, it is our hope that it will stimulate discussion among NATO members and will ultimately lead to improved reliability and serviceability of the human component of the air weapons system.

REFERENCES

1. Macnamara, W.D.
A Survey of Flying Time Regulations,
Unpublished, 1972.

2. Nicholson, A.N.
Duty Hours and Sleep Patterns in Transport Aircrew Operating Long Haul Routes
Flying Personnel Research Committee Report 1302, Ministry of Defence,
London, 1970.

3. Nicholson, A.N.
Military Implications of Sleep Patterns in Transport Aircrew
Proc. Roy. Soc. Med. 1970, 63, 570-572.

4. Hartman, B.O., and H.B. Hall,
Findings on the Cost of Flying Transport Missions
Proceedings of Third Meeting, The Technical Cooperation Program, Technical Panel U-2,
Human Performance and Military Capability, Defence and Civil Institute of Environmental Medicine,
Toronto, Ontario, Canada, 18-20 Sept., 1972.

5. Morgan, B.B. and E.A. Alluisi
Applicability of Research on Sustained Performance, Endurance and Work-Rest
Scheduling to the Development of Concepts and Doctrine of Continuous Operations
Hodge, D.C. Military Requirements for Research on Continuous Operations Technical
Memorandum 12-72, Human Engineering Laboratory, U.S. Army Aberdeen Research and
Development Center, Maryland, 1972.

6. Klein, K.E., H.M. Wegmann, B.I. Hunt
Desynchronization of Body Temperature and Performance
Circadian Rhythm as a Result of Outgoing and Homegoing
Transmeridian Flights
Aerospace Med. 1972, 43, 119-132.

7. Harris, W., and J.F. O'Hanlon
A Study of Recovery Functions in Men
U.S. Army Technical Memorandum 10-72
Human Engineering Laboratory, U.S. Army Aberdeen Research and Development Center, Maryland, 1972.

DISCUSSION

QUESTION	To which category of pilots to the figures on flying hours refer?
MACNAMARA	These figures were obtained from the air forces directly.
QUESTION	I just cannot see how any fighter pilot could fly any more than 40-50 hours. As far as flying instructors anything between 30-45 hours is, I think, about the maximum. In our regulations the maximum hours for instructors are 40 per month on jet aircraft. Perhaps my RAF colleagues would bear me out with the regulations at present in force for the Royal Air Force.
HUTCHINSON	In the RAF the figures for fighter pilots are based on sorties. We allow a maximum of 4 sorties in a day, but there is also an hours maximum which depends on what you have been doing in the sortie. The flying instructors limit is 50 hours in a month, the highest figure I have ever achieved instructing is 90 hours in one month on a light aircraft. I quite agree that is quite debilitating and requires a month's leave to recuperate. We have very strict working hours which require that you have at least 8 hours continuous rest in any 24 hour period and a maximum working period of 16 hours at any one stretch.
BORG	There is a standardisation agreement (STANAG) on this matter for NATO countries.
CHASE	Our US Army helicopter pilots in Veitnam flew over 100 hours a month and very often as high as 125 hours a month.
WILKINSON	My suggestion was that we should permit our commanders to have more flexibility in managing the hours of their crews. I would like to sound a congratulatory note on this attempt to provide a first step towards assessing these problems. Would you try to assess all those factors which you have listed? You would need observers but do you think you would get enough co-operation from the crew to have them fill out their own check lists.
MACNAMARA	In the past one of the difficulties has been that we run into a squadron and as a group of specialists we tell the commander he has a problem when he doesn't have one! The approach we encourage is for the squadron commander to look at the problem with us. In this sense we need to work closely with the crews using their own information.

Endocrine-Metabolic Indices of Aircrew Workload: An Analysis Across Studies

by

*Henry B. Hale, Ph.D.
**Richard C. McNee, M.S.
*James P. Ellis, Jr., Ph.D.
*Ralph R. Bollinger, M.D.
*Bryce O. Hartman, Ph.D.

*Environmental Physiology Branch (VNE) and
**Design and Analysis Branch (BRD)
USAF School of Aerospace Medicine
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas, USA, 78235

SUMMARY

Endocrine-metabolic measures have been used at the USAF School of Aerospace Medicine in a series of field and laboratory studies performed during the past decade. The field studies involved different kinds of military aircraft as well as missions of varying nature and length. Certain aspects were studied in simulated flights conducted under laboratory conditions. The data accumulated in the ten-year period have now been subjected to a cross-sectional analysis in an effort to ascertain the basic relation of endocrine-metabolic activity to the workload in either actual or simulated flights. For the present purpose, load represents degree of flight difficulty multiplied by duration. Difficulty was based upon USAF expert rankings, and duration was based upon fractions of a day. Multiple linear regression analysis was performed on data for urinary epinephrine, norepinephrine, 17-OHCS, urea, Na, K, and the Na/K ratio. This report presents the findings in the first phase of the cross-sectional study. Definition of the utility of endocrine-metabolic assessments of workload in the flight situation is expected to emerge ultimately from additional analyses.

INTRODUCTION

During the past decade the USAF School of Aerospace Medicine has conducted, under either field or laboratory conditions, studies of human responses to flight. The working hypothesis has been that flying operations act in the manner of stressors, eliciting interrelated endocrine-metabolic responses which are compensatory in nature, tending to maintain a state of physiologic balance (homeostasis). A battery of urinary determinations was used to assess the physiologic cost in a wide variety of circumstances, including flying operations of various types and durations which took place at various times of day and utilized a variety of aircraft. Statistical evaluation consistently indicated elevation in physiologic "cost" which apparently related to (a) type of aircraft, (b) flight complexity, (c) flight duration, (d) time of day, and/or (e) crew position. Additional factors also seemed to be contributory, affecting some or all of the endocrine-metabolic functions under study. Preflight circumstances unquestionably exerted recognizable influence on the physiologic state in a subsequent flight period. The data collected in this ten-year effort are now being subjected to a cross-sectional analysis, hoping to bring out basic features which possibly can be useful to flight surgeons, commanders, or aircraft designers. This is a preliminary report of this cross-sectional study.

METHODS

Seventeen of the previously-published stress studies (1-13) and two unpublished studies provided the data which were to be considered from a spectral approach. As the first step in this collective study, multiple linear regression analysis was performed. The "load" for each of the studies was computed in each case using a value taken from a 7-point "complexity" scale multiplied by a value from a 5-point duration scale. The complexity (degree of difficulty) scale used expert ratings on the difficulty of flying the several different aircraft in our studies which were supplied by the USAF Directorate of Aerospace Safety, Norton AFB, California. The duration scale used 4-hour increments based upon a 24-hour period. The "loads" thus established comprised a spectrum which was regarded tentative, and the physiologic trends were used to predict loads. Thus, for these man-machine combinations, man is considered a "sensor."

The urinary determinations are as follows: (a) epinephrine, an index of adrenomedullary activity, (b) norepinephrine, an index of sympathetic nervous system activity, (c) 17-hydroxycorticosteroids (17-OHCS), an index of adrenocortical activity, (d) urea, an index of protein catabolism, and (e) sodium and potassium, indices of mineral metabolism. The ratio of sodium to potassium is considered an index of metabolic balance (homeostasis).

As background, it is essential to know the character of the circumstances of varying load. Only endflight data were obtainable in certain of the studies. In some of these cases the flights terminated in daytime; in others, there was nighttime termination. Where within-flight determinations were made, the overall average was used. If the entire flight (whether real or simulated) took place in daytime, the average of the within-flight values was placed with the endflight values for daytime. Similarly, averaged within-flight data for flights made in nighttime periods were placed with endflight values obtained at night. In very prolonged flights the averaged within-flight data were placed in a separate category, namely, combined day-night.

Daytime data were obtained in the following cases: (a) 18-hour F-4C flights, (b) B-52 test flights lasting 9-14 hours, (c) 8-hour FB-111 test flights, (d) C-135B staged global flights lasting 6 days,

(e) 6-hour F-104 transatlantic flights, (f) staged F-102 transoceanic flights, (g) 6-hour F-100 transoceanic flights, (h) a 12-hour simulated flight, and (i) a 4-hour period of normobaric hyperoxia.

The combined day-night studies included (a) a 28-hour HH-3E transatlantic flight, (b) 54-hour double-crew C-141 intercontinental flights, (c) transoceanic C-135B flights made during a 7-week airlift exercise, (d) 66-hour double-crew C-5 intercontinental flights, (e) C-141 staged intercontinental flights of 5-7 days duration, (f) 2-hour F-100 training flights, and (g) air traffic control work.

Nighttime termination studies were limited to the following: (a) 20-hour C-130E round-trip New Zealand-Antarctica flights, (b) a 36-hour simulated flight, and (c) night shift work.

RESULTS

The endocrine-metabolic values for the individual studies are given in Tables I to III. In each table the studies are arranged in descending order with respect to load. None of the endocrine-metabolic indices shows clear-cut relationship to load. Obviously, task complexity, duration, and time of day are not the only factors affecting physiologic cost. It is reasonable to consider the assigned loads faulty, thereby contributing to what appears to be nonpatterned variation. Table IV shows the results of the multiple linear regression analysis. On the basis of the data from all 19 studies, combinations of endocrine-metabolic variables other than the catecholamines showed statistically-significant load-dependence when adjusted by use of a numerical factor for time of day (night = 0, day-night = 1, and day = 2). The endocrine-metabolic variables which, collectively, showed the best relationship to load were sodium, potassium, and the sodium-potassium ratio. As shown in Table IV and Fig. 1, load can be roughly predicted by this complex of variables ($r = .81$). The regression equation appears as a footnote to the table.

Since epinephrine and norepinephrine data were unobtainable in two of the studies, a secondary analysis was performed on the 17 studies which had complete data. Potassium, Na/K, norepinephrine, and time of day collectively gave the most clear-cut load-dependency ($r = .84$) (Table IV and Fig. 2). Using more of the physiologic variables did not greatly improve this prediction.

DISCUSSION

As we indicated in the introduction to this paper, our working hypothesis is that the task of flying involves a complex of stressors acting in some unspecified combination to produce changes in the physiological status of the crew. We interpret our findings and make recommendations regarding specified problems in flying operations within this framework. However, this step from a physiological change to an operational recommendation is indeed a large one, perhaps too large to withstand close scrutiny. We have, therefore, undertaken a critical analysis of this entire approach, using our own studies as a starting point.

Workload is a more concrete variable against which to perform our first analysis than our working hypothesis implies, but it is obviously of great interest to operational groups, so we accepted the challenge. Many procedural and technical problems surfaced once we began rating workload and identifying the appropriate endocrine-metabolic data to be evaluated, and compromises were made which may have weakened the analysis from a physiological (not statistical) point of view. In addition, there are three major aspects of the study which are of concern. First, the workload rating procedure is less than completely satisfactory, in that it does not take into account operationally-significant factors which logically deserve consideration, such as the type of mission, the tactical and/or environmental threat, etc. The outcome of the multiple linear regression analysis is no better than the workload ratings which were assigned. It does, however, have the advantage of being objective rather than subjective. Second, the workload rating procedure resulted in many ties and considerable clumping in the middle of the range of scores. This can compromise the efficiency of the statistical analysis. Third, the statistical demands of the analysis would not fully accommodate all of the physiology which deserves consideration. In particular, we were unable to consider the highly useful time courses of the physiological changes without discarding so many of the early studies that the analysis itself would be compromised.

Nevertheless, it was the impartial assessment and judgment of the statistician to whom we chose to submit our case. The results from this first analysis are less than a complete indorsement of our approach and more than the bitter disappointment of random fluctuation. We consider them sufficiently promising to proceed with further analysis for this sub-evaluation.

REFERENCES

1. Marchbanks, V. H., Jr., H. B. Hale, and J. P. Ellis, Jr. Stress responses of pilots flying 6-hour overwater missions in F-100 or F-104 aircraft. Aerospace Med. Vol. 34, 1963, pp 15-18.

2. Hale, H. B., E. W. Williams, J. E. Anderson, and J. P. Ellis, Jr. Endocrine and metabolic effects of short-duration hyperoxia. Aerospace Med. Vol. 35, 1964, pp 449-451.

3. Hale, H. B., J. C. Duffy, J. P. Ellis, Jr., and E. W. Williams. Flying stress in relation to flying proficiency. Aerospace Med. Vol. 36, 1965, pp 112-116.

4. Hale, H. B., J. P. Ellis, Jr., and E. W. Williams. Endocrine and metabolic changes during a 12-hour simulated flight. Aerospace Med. Vol. 36, 1965, pp 717-719.

5. Kramer, E. F., H. B. Hale, and E. W. Williams. Physiological effects of an 18-hour flight in F-4C aircraft. Aerospace Med. Vol. 37, 1966, pp 1095-1098.

6. Hale, H. B., C. A. Anderson, E. W. Williams, and E. Tanne. Endocrine-metabolic effects of unusually long or frequent missions in C-130E or C-135B aircraft. Aerospace Med. Vol. 39, 1968, pp 561-570.

7. Demos, G. T., H. B. Hale, and E. W. Williams. Anticipatory stress and flight stress in F-102 pilots. Aerospace Med. Vol. 40, 1969, pp 385-388.

8. Hale, H. B., E. W. Williams, and C. J. Buckley. Aeromedical aspects of the first nonstop transatlantic helicopter flight. III. Endocrine-metabolic effects. Aerospace Med. Vol. 40, 1969, pp 718-723.

9. Hale, H. B., E. W. Williams, B. N. Smith, and C. E. Melton. Excretion patterns of air traffic controllers. Aerospace Med. Vol. 42, 1971, pp 127-138.

10. Hale, H. B., E. W. Williams, B. N. Smith, and C. E. Melton. Neuroendocrine and metabolic responses to intermittent night shift work. Aerospace Med. Vol. 42, 1971, pp 156-162.

11. Hale, H. B., B. O. Hartman, D. A. Harris, E. W. Williams, R. E. Miranda, J. M. Hosenfeld, and B. N. Smith. Physiologic stress during 50-hour double-crew missions in C-141 aircraft. Aerospace Med. Vol. 43, 1972, pp 293-299.

12. Hale, H. B., W. F. Storm, J. W. Goldzieher, B. O. Hartman, R. E. Miranda, and J. M. Hosenfeld. Physiological cost in 36- and 48-hour simulated flights. Aerospace Med. Vol. 44, 1973, pp 871-881.

13. Hale, H. B., B. O. Hartman, D. A. Harris, R. E. Miranda, and E. W. Williams. Physiologic cost of prolonged double-crew flights in C-5 aircraft. Aerospace Med. Vol. 44, 1973, pp 999-1008.

TABLE I

DAYTIME STUDIES

Test Circumstance	Load	Endocrine-Metabolic Index						
		E	NE	17-OHCS	U	Na	K	Na/K
F4C flight	30	0.74	3.80	444	1.39	6.1	1.8	4.0
B52 flight	18	1.52	5.09	691	1.69	12.5	4.6	2.7
FB111 flight	18	0.83	5.55	395	1.34	15.0	4.7	3.4
C135B flight	15	--	--	338	1.16	9.4	3.6	2.8
F104 flight	14	1.83	5.72	298	1.26	14.6	4.2	3.6
F102 flight	14	1.60	4.33	341	1.78	11.1	3.9	2.9
F100 flight	14	1.28	3.57	373	1.52	11.2	4.5	2.6
12-hr Simulated flight	6	1.56	4.92	462	1.82	12.4	3.9	3.8
Hyperoxia	1	0.56	2.88	325	1.33	11.4	5.1	2.6

E, NE, and 17-OHCS = epinephrine, norepinephrine, and corticosteroids (μg/100 mg creatinine); U = urea (gm/100 mg creatinine); Na and K = sodium and potassium (mEq/100 mg creatinine); and Na/K = mean ratio of sodium and potassium, not ratio of mean sodium and mean potassium.

TABLE II

DAY-NIGHT STUDIES

Test Circumstance	Load	Endocrine-Metabolic Index						
		E	NE	17-OHCS	U	Na	K	Na/K
HH3E flight	26	1.70	3.93	217	1.40	7.1	2.1	3.4
C141 double-crew flight	15	1.11	4.67	294	1.32	12.4	3.2	4.3
C135 flight	10	1.16	3.37	369	1.47	10.6	2.6	4.6
C5 flight	10	1.07	4.83	305	1.24	11.4	4.0	3.3
C141 single-crew flight	8	0.99	4.30	212	1.30	10.7	3.1	4.0
F100 training flight	7	0.74	2.69	280	1.16	12.4	3.6	3.8
Air traffic control work	6	1.60	4.75	289	1.35	11.7	3.8	3.5

See Table 1 for explanation of terms.

TABLE III

NIGHTTIME STUDIES

Test Circumstance	Load	Endocrine-Metabolic Index						
		E	NE	17-OHCS	U	Na	K	Na/K
C130E flight	20	0.99	3.63	196	1.22	7.5	2.4	3.2
36-hr Simulated flight	9	--	--	243	1.35	10.3	2.4	4.6
Sedentary work	2	0.68	3.18	274	1.30	9.2	2.8	3.4

See Table 1 for explanation of terms.

TABLE IV

ASSIGNED AND PREDICTED LOADS

Study	Assigned Load	Predicted Load*	Predicted Load**
(1) F4C	30	27	29
(2) HH3E	26	24	24
(3) C130E	20	15	16
(4) B52	18	11	11
(5) FB111	18	11	11
(6) C135B, global	15	19	--
(7) C141, double-crew	15	10	10
(8) F102	14	18	18
(9) F104	14	17	15
(10) F100	14	12	11
(11) C135B, transoceanic	10	12	9
(12) C5	10	6	10
(13) 36-hr simulated flight	9	7	--
(14) C-141, staged	8	11	12
(15) F100, training	7	10	3
(16) 12-hr simulated flight	6	12	14
(17) Air traffic control work	6	7	10
(18) Night shift duty	2	11	9
(19) Hyperoxia	1	2	3

*Load = 71.3 + 3.35(Na) + 7.42(time) - 17.9(K) - 12.1(Na/K)
**Load = 46.8 + 3.70(NE) + 6.16(time) - 9.70(K) - 6.70(Na/K)

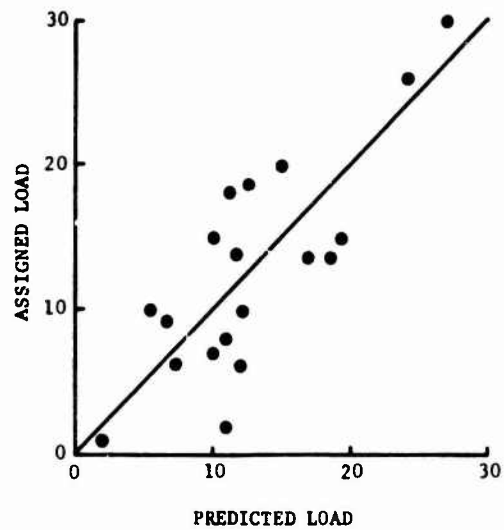


Fig. 1. Predicted load versus assigned load

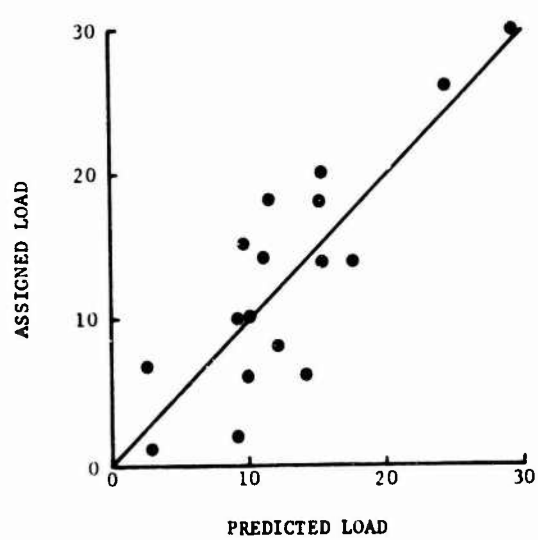


Fig. 2. Predicted load versus assigned load

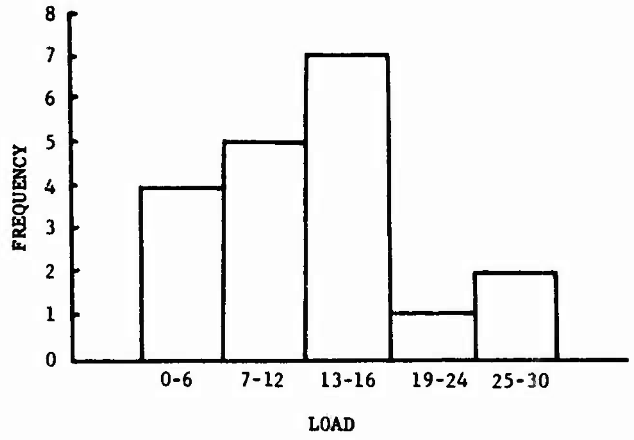


Fig. 3. Load distribution

TIME DEPENDENCE OF THE FLIGHT INDUCED INCREASE OF FREE URINARY CORTISOL SECRETION IN JET PILOTS

G. Ulbrecht, Col., GAF, MD; E. Meier, Cpt., GAF,MD; R.Rothenfußer, GAF, Research Technician, K.v.Werder,Cpt.,GAF,MD
German Air Force Institute of Aviation Medicine
808 Fürstenfeldbruck, Germany

SUMMARY

A modified competitive protein binding assay of free urinary cortisol using a single solvent extraction and a cortisol binding globulin from a dexamethasone suppressed male subject was developed. The seperation of bound and free cortisol was performed by adsorption of the free cortisol to dextran coated charcoal. The sensitivity of the method allows to measure as low as 0.2 ng per tube. The coefficient of variation within one assay is only 4.8 %, which makes this method suitable to measure minute changes of free urinary cortisol excretion during the day in a fractionated 24 hr-urine collection. The mean free cortisol excretion in 35 normal men was 63 ± 3 µg/24 hrs ± SE showing the expected circadian rythm of adrenocortical activity. In seven F-104 pilots flying two missions a day the 24 hr free urinary cortisol secretion was significantly higher (89 ± 12 µg/24 hrs ± SE) compared to 12 pilots on day of rest (43 ± 7 µg/24 hrs.). When 26 F-104 pilots, 12 RF 4E pilots and 14 weapon system operators (WSO) were evaluated by measuring free urinary cortisol excretion in short intervals it could be demonstrated, that only the pilots flying early in the morning showed an enhancement of adrenocortical activity compared to normal controls, suggesting a change of excitability of the hypothalamo-pituitary-adrenal system during the day. This might have to be taken into account if investigations about stress of flying and adrenocortical activity are carried out.

INTRODUCTION

It has been shown that there is a variety of different stressful situations leading to the activation of the human adrenal cortex (1). The increase of cortisol secretion is always induced by an increase of circulating adrenocorticotropic hormone (ACTH) (2). The ACTH-secretion from the anterior pituitary is again stimulated by the release of corticotropin releasing factor (CRF) from the hypothalamus (3) which is under control of higher brain centers (4) mediating the release of CRF (Fig. 1).

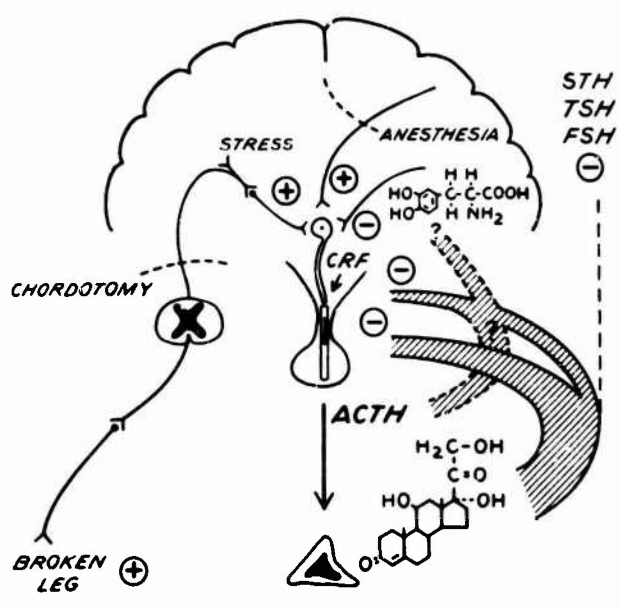


Fig. 1: The hypothalamo-pituitary-adrenal system (CRF → ACTH → Cortisol). Stimulating - ⊕ and inhibiting - ⊖ factors upon CRF - secretion are depicted. Stress leads only to CRF-secretion if the neural transmission is not interrupted (chordotomy). CRF-secretion is also influenced by general anesthesia (5) and biogenic amines (6). The shaded pathways represent the negative feedback systems which exist between the adrenal and the pituitary and hypothalamus (long feedback), the pituitary and the hypothalamus (short feedback) and probably between the pituitary and higher brain centers (4). There is also a negative unspecific feedback of cortisol upon other pituitary hormones (somatotropic hormone = STH, thyrotropin = TSH, and follicle stimulating hormone = FSH).

The magnitude of the increase of the cortisol plasma level reflects the functional state of the hypothalamo-pituitary-adrenal axis as well as the intensity of the stress activating this system. Since flying a jet is a situation in which the pilot has to cope with a variety of psychological and biophysical stresses it is not surprising that adrenocortical activity in pilots is found to be enhanced on day of flight (7,8). If the activation of the hypothalamo-pituitary-adrenal-system is due to psychological influences or to the actual physical stress has not been clarified yet. Since it is difficult to measure ACTH or cortisol blood levels in a group of pilots during the actual time of flying one has to rely on the measurements of cortisol or cortisol metabolites in the urine. In most of the previous investigations concerning the influence of flying upon adreno-cortical activity the conjugated 17-hydroxycorticosteroids were measured(,7,8). Since the nonmetabolized socalled free urinary cortisol is in contrast to the conjugated 17-OHCS independent of the cortisol binding protein moieties in the serum and of liver function, it reflects far better than these 17-OHCS the instantaneous functional state of the adrenal cortex (9,10). This has been proven in clinical medicine for the diagnosis of adrenocortical abnormalities. We have developed a simple technique for the measurement of free urinary cortisol which is based on the competitive protein binding radio-ligand-assay originally developed for the determination of plasma corticoids by MURPHY and PATTEE (11) and modified for the measursment of urinary cortisol (12).

MATERIALS AND METHODS

The basic principles of the competitive protein binding analysis (CPBA) has been discussed previously (13). Our method employed is patterned closely after that described by HSU and BLEDSOE (12). It differs in that the assay system employs a ^{3}H-cortisol labeled cortisol-binding-globulin (CBG) solution and dextran-coated charcoal adsorbent. Only one aliquot of the dichloromethane extract of the urine is subjected to a single alkaline wash before the CPBA is performed.

1, 2, 6, 7 -^{3}H-Cortisol with a specific activity of 290 mCi/mg was purchased from the Radiochemical Centre, Amersham, U.K.. Highly purified dichloromethane was obtained from E. Merck, Darmstadt, Germany, Dextran T 70 (Pharmacia, Uppsala, Sweden) and Norit A charcoal (Serva, Heidelberg, Germany) were used for the adsorbent. The CBG-serum was obtained from a normal male subject 4 hours after oral administration of 2 mg dexamethasone. Liquid scintillation counting was performed in a Packard Tri-Carb using Insta-Gel (Packard Inst.Comp., Downers Grove, Ill., Cat.No. 6002174) as scintillator system.

Cortisol standard curves ranged conventionally from 10 to 100 ng/m. To 0.1 ml of the standard solution 0,7 ml of 0.05 m phosphate buffer, pH 7.4 was added. The CBG-serum was diluted 1:50 in phosphate buffer and an equal volume of the ^{3}H-cortisol solution (also diluted in phosphate buffer) with 40.000 cpm/ml was added (Fig. 2). 0.2 ml of this CBG - ^{3}H-cortisol- solution was pipeted into the incubation mixture, after which an incubation at 40° C for 5 minutes was followed by 30 minutes at 4° C.

Dextran-coated charcoal was prepared by mixing equal volumes of 0.5% Norit A charcoal in phosphate buffer with 0.05% Dextran 10 in phosphate buffer. This solution was kept in an ice bath and constantly stirred when used. 0.5 ml of this solution was added to the incubation mixture (Fig. 2) after which all samples were shaken vigorously and kept in an ice bath for 10 minutes. After centrifugation 0.5 ml of the supernatant was pipeted into a counting vial together with 10 ml of Insta-Gel and counted for 10 minutes.

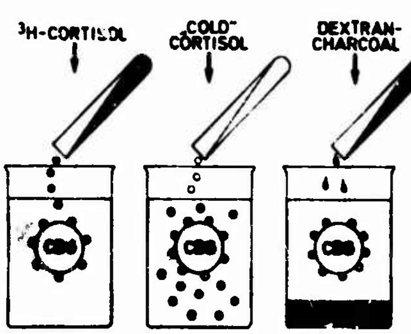


Fig.2: Schematic picture of the CPBA for cortisol using dextran-coated charcoal as adsorbent.
1. The CBG-^{3}H-cortisol solution is prepared (left).
2. An equilibrium between free and CBG-bound labeled and unlabeled cortisol is established (center).
3. The free labeled and unlabeled cortisol is bound to the charcoal and the CBG-bound cortisol is in the supernatant (right).

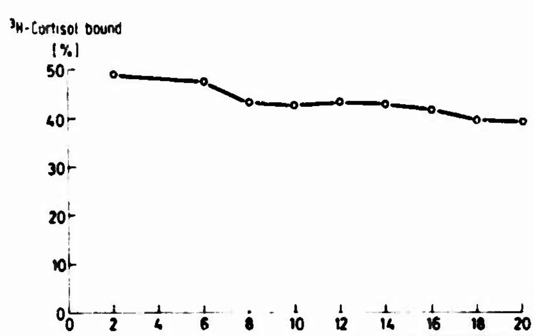


Fig.3: Influence of time after charcoal addition on ^{3}H-cortisol binding to CBG. The adsorption of ^{3}H-cortisol to charcoal from 8 to 14 minutes after adding the charcoal solution is not significantly different as is demonstrated by the constant ^{3}H-cortisol binding to CBG. This plateau leaves enough time to handle up to 60 samples. After 14 minutes a fall of ^{3}H-cortisol binding to CBG is observed due to dissociation of ^{3}H-cortisol from its CBG-binding sites.

A 1/1000 th or 1/2000 th aliquot of a 24 hr-urine collection is extracted with 10 ml of dichloromethane. The aqueous layer is discarded by aspiration and the organic phase is washed once with 0.1 normal sodiumhydroxide. 5 ml of the dichloromethane extract is dried after which 0.8 ml phosphate buffer and 0.2 ml of the ^{3}H-cortisol-CBG solution is added. This is incubated and handled in the same fashion as the standard curve. A control urine is measured with each assay. The coefficient of variation within one assay is 4.8%, the variation from day to day does not exceed 13%. The crossreactivity with other steroids is shown in fig. 4.

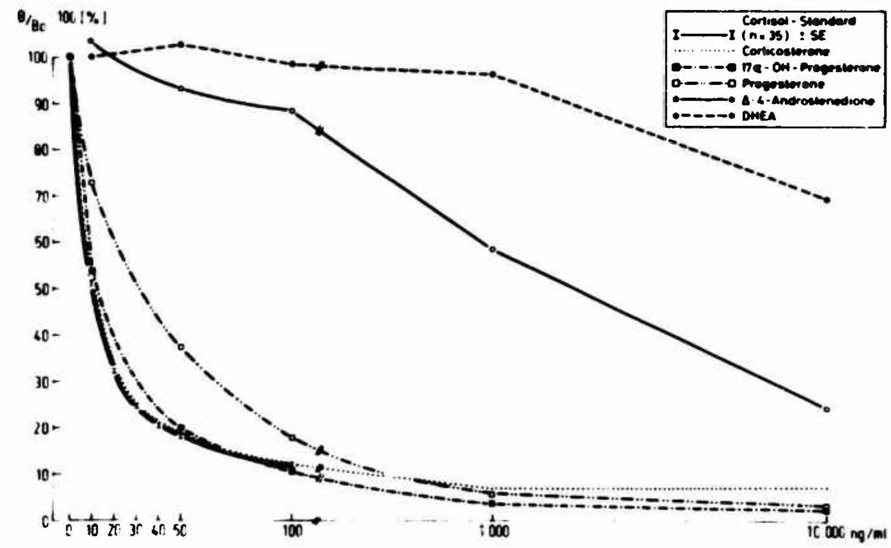


Fig. 4: Specificity of Steroid-Binding to CBG. Corticosterone and 17-OH-progesterone show complete crossreactivity. Progesterone shows partial crossreactivity whereas Δ - 4 - androstenedione and dehydroepiandrosterone (DHEA) do not significantly interfere with the assay. Since there are only minute quantities of free corticosterone, 17-OH progesterone, and progesterone in the urine compared to cortisol, the crossreactivity of these steroids can be neglected.

With this method free urinary cortisol excretion in 24 hr-urine was measured in 35 normal males (physicians), 12 jet-pilots on day of rest and 7 pilots flying two missions a day. The circadian rythm of cortisol secretion was evaluated in 15 normal males measuring free urinary cortisol every 3 hours over day-time and in one nighttime collection (6 hours). These results were compared with the free urinary cortisol excretion of 26 F-104 pilots, 12 RF-4E-pilots and 14 weapon system operators (WSO). The pilots and WSOs emptied their bladder after they got up in the morning. The time was noted and the urine discarded. Before the flight they emptied their bladder again. This urine was collected and the time noted. Roughly 3 hours later another urine was collected. During this collection period the pilots flew a training mission (high and low level, range, ACT) lasting from 50 to 120 minutes. The second urine collection was followed by a third one, again covering a period about 3 hours. Free urinary cortisol excretion was calculated in ng/min. Since the collection period varied the first urine had to be regarded as baseline urine, from which changes in free urinary cortisol excretion in the following urines were calculated and compared with the normal circadian rythm of free urinary cortisol excretion. Therefore all data had to be plotted in percent increase or decrease from basal cortisol excretion.

RESULTS

The mean free urinary cortisol excretion in 35 normal males (resident doctors of a university medical center) was 63 ± 3 μg/24 hr ± SE with a normal range from 27 to 99 μg/24 hrs (± 25). Five patients with adrenal insufficiency had levels ranging from less than 5 to 24 μg/24 hrs and two patients with Cushing's syndrome had a free urinary cortisol excretion of 846 μg/24 hrs, respectively 360 μg/24 hrs. Seven jet pilots (F-104) flying two missions a day excreted 89 ± 12 μg/24 hrs ± SE compared to 43 ± 7 μg/24 hr of 12 pilots on day of rest.

In 15 normal subjects the typical circadian rhythm of free urinary cortisol excretion could be established (Fig. 5).

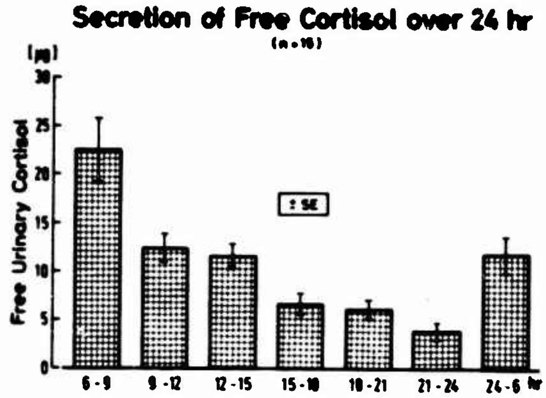


Fig.5: Free urinary cortisol excretion in 15 normal subjects.
Free urinary cortisol was measured 6 a.m. to midnight every 3 hours and in one night time urine collection (6 hrs)

The free urinary cortisol secretion pattern in 26 F-104 pilots flying a mission between 6 a.m. and 3 p.m. is depicted in Fig. 6.

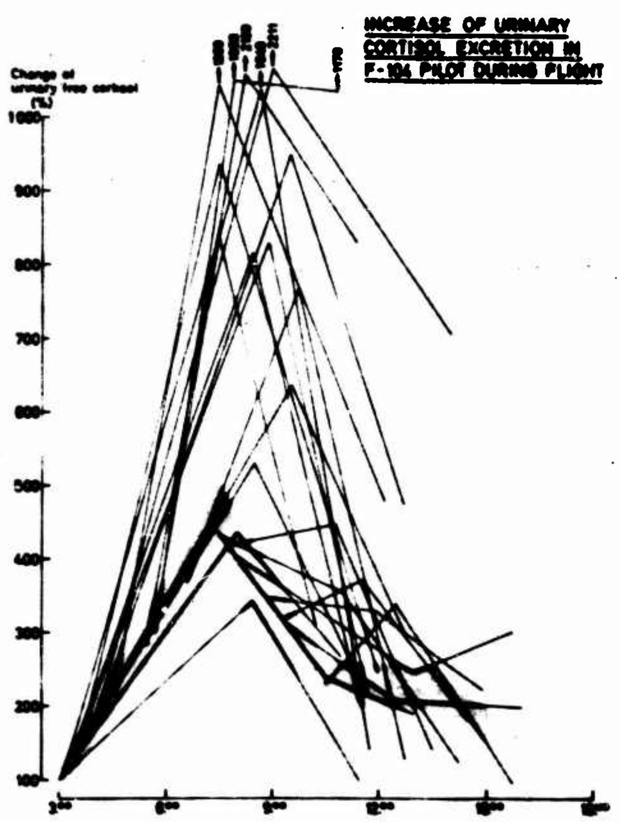


Fig.6: Changes of free urinary cortisol excretion during flight.
The mission (50-120 minutes) took place between 6 a.m. and 3 p.m. The change was calculated in percent from the preflight excretion and compared with tha normal circadian periodicity (shadowed area = mean ± SE).

It waa apparent from these data that only thoae pilots had exaggerated increases in morning urinary cortisol excretion who flew early in the morning, whereas the othera showed no significant deviation in their cortisol excretion from the normal pattern. When free urinary cortisol excretion was measured in 12 additional RF-4E-pilota and 14 weapon system operators we plotted these data according to tha commencement of their mission. Arbitrarily the pilots starting their mission bafore 9 a.m. were put togethar in ona group and were compared with those whoae mission had begun paat 9 a.m. (Fig. 7). A striking difference could be observed in that moat of the pilota flying aarly in tha morning ahowed an increase of their cortisol axcretion compared to nonflying individuala whereas the pilots flying latar had in most instances a complataly normal circadianrhythm of adrenocortical activity.

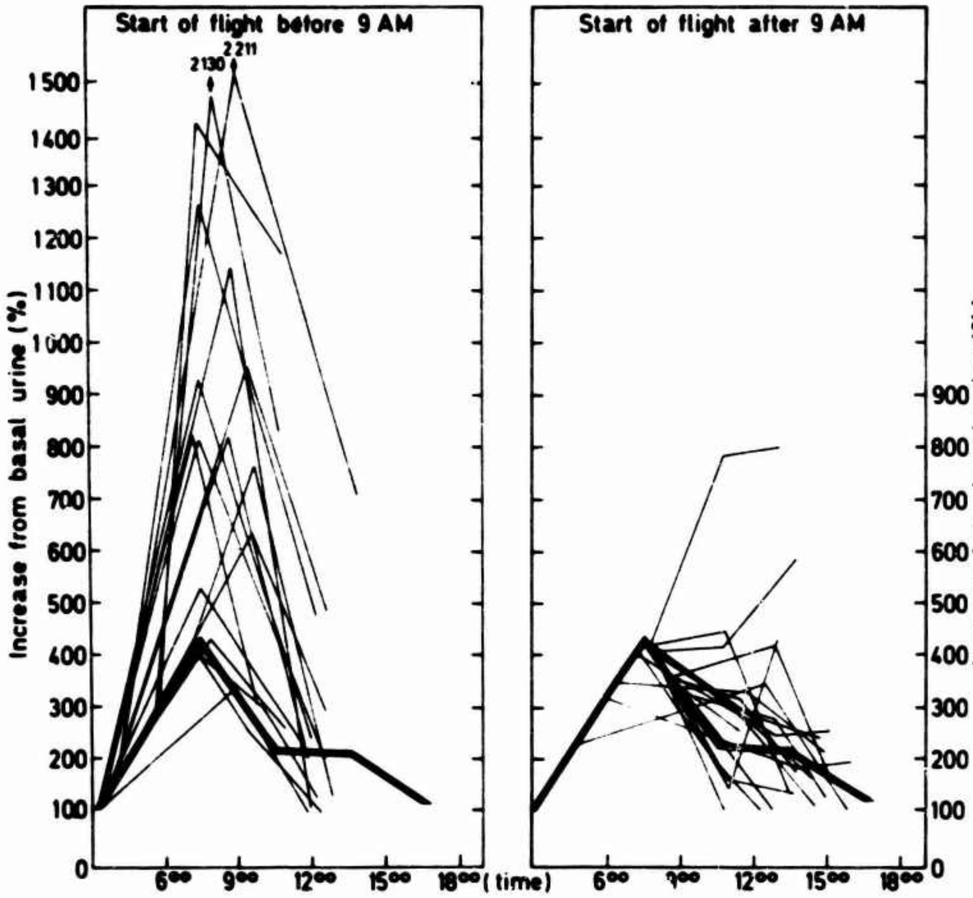


Fig.7: Change of free urinary cortisol excretion during flight in jet pilots and weapon system operators. The calculation was performed as in fig. 6. Whereas most of the pilots flying before 9 a.m. show increases in cortisol excretion compared to the nonflying group (thick line), the pilots starting with their missions past 9 a.m. show generally a normal circadian rhythm of adrenocortical activity.

DISCUSSION

To evaluate adrenocortical activity in pilots during flight one has to rely usually on the measurement of cortisol or cortisol metabolites in the urine since obtaining plasma samples at short intervals under inflight conditions is impracticable. Most of the investigators studying inflight adrenocortical activity estimated the excretion of cortisol metabolites by the available standard procedures (7,8). The half life time of plasma cortisol is about 80 minutes (5,14) and most of the secreted cortisol is reduced to the tetrahydroform and coupled to glucuronic acid in the liver. The measurement of the conjugated 17-hydroxycorticosteroids (17 - OHCS) therefore does not reflect the instantaneous state of cortisol secretion since the overall 17-OHCS excretion depends also on liver function which might be impaired during jet flight due to changes in hepatic circulation. The free urinary cortisol excretion is independent of liver function and reflects far better the actual state of adrenal activity (9), since only the free plasma cortisol, i.e. the biologically active fraction which is not bound to the cortisol binding globulin (CBG) or albumin is found in the urine. Since binding capacity of CBG is limited and the affinity for cortisol binding of albumin is low (15), an increase of total plasma cortisol levels is accompanied by an even greater increase of the free, unbound cortisol, which will be reflected by the urinary free cortisol excretion, thus serving as an amplifier for the detection of adrenocortical activation. The previous methods for the measurement of free urinary cortisol were based on colorimetry and fluorometry necessitating chromatography (16). With the advent of the CPBA for steroids this has changed. Our method involving a solvent extraction of the urine with dichloromethane, a single alkaline wash and subsequent CPBA with a ^3H-cortisol-CBG solution fulfills all criteria for a practicable laboratory procedure. The sensitivity, precision and reproducibility is good, the method is fast and only small quantities of urine are needed. Our results of free urinary cortisol excretion in normals and in patients with adrenal abnormalities are in good agreement with other investigators (12).

Enhancement of adrenocortical activity in jet pilots during flight has been shown previously (7,8). Wether this is due to actual physical stress or due to psychological factors is not clear yet (17) Activation of the hypothalamo-pituitary-adrenal system represents the common terminal pathway of the answer to a variable and sometimes complex stressor. Our finding that pilots flying missions early in the morning seem to have a higher cortisol secretion compared to those who start their flight in the late morning hours suggest a rythm of hypothalamic-pituitary adrenal responsiveness. Such cyclicity of the responsiveness of hypothalamo-pituitary adrenal-system to stressors has been reported and discussed before (18,19). If our finding can be explained only by time factor can not be stated with absolute assurance since, though the training of all pilots studied was comparable, two different jets were flown. It has been shown that this alone can cause different responses regarding adrenal activation (20). Though absolute values of cortisol excretion would have helped clarifying this phenomenon, because of the irregular urine collection periods due to the pilot's training program it was not possible to compare the absolute values obtained in our study.

Taking all this in account the difference in the response of inflight adrenocortical activation depending on the daytime of flight and most likely the state of the endogenous circadian rythm of cortisol secretion, is striking. It has to be further investigated before any conclusions should be drawn but should be looked for if studies concerning adrenocortical activity in response to complex stressors are carried out.

REFERENCES

1. Yalow, R.S., N. Varsano-Aharon, E. Echemendia, S.A. Berson: HGH and ACTH secretory responses to stress. Horm. Metab. Res. 1 (1969) 3
2. Scriba, P.C., K.von Werder, K.Schwarz: Hypothalamus und Hypophyse.In: W.Siegenthaler:Klinische Pathophysiologie, Thieve Verlag, Stuttgart, 1973
3. Vernikos-Danellis, J.: Estimation of Corticotropin-releasing activity of rat hypothalamus and neurohypophysis before and after stress. Endocrinology 75 (1964) 514.
4. Martini, L., F. Fraschini, M. Motta: Neural control of anterior pituitary function. Rec.Progr. Horm. Res. 24(1968) 439.
5. Werder, K.von, W.C. Stevens, T.H. Cromwell, E.I. Eger, P.H. Forsham: Adrenal function during long-term anesthesia in man. Proc.Soc. Exper. Biol.Med. 135(1970) 854.
6. Fuxe, K., T. Hökfeld: Catecholamines in the hypothalamus and the pituitary gland. In: Frontiers in Neuroendocrinology, E.: W.F. Ganong and L. Martini, Oxford Univ.Press, New York, 1969.
7. Marchbanks, V.H.: Effect of flying stress on urinary 17-hydroxycorticosteroid levels. J.Aviat. Med. 29 (1960) 639.
8. Colehour, J.K., A. Graybiel: Excretion of 17-hydroxycorticosteroids, catecholamines and uropepsin in the urine of normal persons and deaf subjects with bilateral vestibular defects following acrobatic flight stress. Aerospace Med. 35 (1964) 370.
9. Forsham, P.H.: The adrenal cortex. In: R.H. Williams, Textbook of Endocrinology, W.B. Saunders Comp., Philadelphia, 1968.
10. Labhart, A.: Klinik der Inneren Sekretion, Springer Verlag, Berlin, Heidelberg, New York, 1971, p. 392.
11. Murphy, B.E.P., C.J. Pattee: Determination of plasma corticoids by competitive protein binding analysis using gel filtration. J.Clin. Endocr. 24 (1964) 919.
12. Hsu, T.H., T. Bledsoe:Measurement of urinary free corticoids by competitive protein-binding radioassay in hypoadrenal states. J.Clin. Endocr. 30 (1970) 443.
13. Murphy, B.E.P.: Hormone assay using binding proteins in blood. In: Principles of competitive protein-binding assays, Ed.: W.D. Odell and W.H. Daughaday, Lippincott Comp., Washington,1971.
14. Milby, J.C., W.W. Spink: Comparative studies on adrenocortical function and cortisol metabolism in healthy adults and in patients with shock due to infection. J.Clin.Invest.31 (1958) 1791.
15. Daughaday, W.H. : Steroid protein interactions. Phys.Rev. 39 (1959) 885.
16. Rosner, J.M., J.C.Cos, E.G. Biglieri, S. Hane, P.H. Forsham: Determination of urinary unconjugated cortisol by glass fiber chromatography in the diagnosis of Cushing's syndrome.J.Clin. Endocr. 23 (1963) 820.
17. Mason, J.W.: A review of psychoendocrine research on the pituitary-adrenal corticol system. J.Psychosomatic Med. 30 (1968) 576.
18. Ichikawa, Y., M. Nishikai, M. Kawagoe, K. Yoshida, M. Homma: Plasma corticotropin, cortisol and growth hormone responses to hypoglycemia in the morning and evening. J.Clin.Endocr.34(1972) 895.
19. Nakgawa, K.: Probable sleep dependency of diurnal rythm of plasma growth hormone response to hypoglycemia. J.Clin. Endocr. 33 (1971) 854.
20. Marchbanks, V.H., H.B. Hale, J.P. Ellis: Stress responses of pilots flying 6-hour overwater missions in F-100 and F-104 aircraft. Aerospace Med. 34 (1963) 15.

Mailingaddress:

Priv.Doz. Dr. G. Ulbrecht, Oberstarzt, Flugmedizinisches Institut der Luftwaffe, 808 Fürstenfeldbruck, Fliegerhorst, BRD.

EMOTIONAL AND BIOCHEMICAL EFFECTS OF HIGH WORK-LOAD

LCDR William B. McHugh, MC USN
Navy Medical Neuropsychiatric Research Unit
San Diego, California 92152

Dr. C. A. Brictson
Dunlap and Associates, Inc.
La Jolla, California

and Dr. Paul Naitoh
Navy Medical Neuropsychiatric Research Unit
San Diego, California 92152

SUMMARY

A preliminary longitudinal multifactorial study of the interrelationships of biochemical, mood, biographical factors and landing performance under high work load conditions has been carried out with U. S. Naval Aviators. Levels of serum cholesterol, serum uric acid, blood lactate, pyruvate, and mood assessments were made during periods of non-flying activity and during periods of increased cumulative work load. Uric acid values fell during moderate cumulative work load, and cholesterol values fell during high cumulative work load. Increased variability of pyruvate and lactate were noted with increased cumulative work load. Increased cumulative work load did not affect emotions or performance but altered mood association patterns and altered the relationships of mood and performance. Experience was correlated with performance under zero cumulative work load conditions. Emotion correlated with performance under high cumulative work load conditions.

INTRODUCTION

The concept of homeostatic control of the internal environment has been a keystone of physiologic theory for many years and many studies of environmental stress have utilized this concept. Until recently however, there have been few comprehensive theories regarding emotional and behavioral aspects of stress response. Teichman has elaborated the concept that the organism is a complex self regulating system that responds to both physical and symbolic stimuli by compensatory responses whose purpose is to control physiologic and behavioral phenomena within certain limits. Thus a stress reaction is defined as a variation in behavior beyond normal limits, and stimuli associated with stress reactions are defined as stressors. In this analysis it becomes appropriate to consider the stressor, the compensatory action of the regulating system, and the phenomena which are controlled (1).

As a stressor is encountered, the stress reaction is first noted in the variation of the compensatory mechanisms, and if these fail then variation exceeding the normal limits would be noted in the response the controlled phenomena. In considering the reaction of an individual to a stressor, only the changes relation to the normal limits of the controlling variables and controlled phenomena need be considered. the reaction of a group of individuals to a stressor, not only must the levels of the controlling va and controlled phenomena be studied, but the heterogeneity of the group's reaction to the stressor must considered. As the strength of the various compensatory mechanisms must vary from one individual to an r so the reactions of the individuals to a uniform stressor will vary and the heterogeneity of the grou reaction as measured by the variance will be increased or decreased. This change, however, must be tempered by the fact that as a variable reaches its biologic limits its range of variations is necessarily restricted. Hence both increases and decreases in variance may be significant. According to Gellhorn, and others, the direction of any compensatory physiological activity induced by a stimulus depends upon the concurrent physical status of the body including the degree of physiologic arousal and the awareness of the stimulus (2). These factors interact with each other and with learning and memory in a complex manner to influence not only the direction but the degree of physiological responses.

Field and laboratory studies indicate that the factors influencing physiologic responses are quite complex. Acute rises in serum cholesterol have been noted in medical students taking examinations and in accountants during periods of increased demand for auditing services (3, 4). Similar rises in serum cholesterol have been noted in UDT trainees during periods of overwhelming demands (5). Indeed, failure to respond biochemically to the stress was a valid predictor of impending failure in the training program. However, adaptation to the demands led to a decline in cholesterol levels (6)

Emotional factors can moderate biochemical responses. In a study of laborers facing iminent unemployment, elevations of serum cholesterol occurred only in those whose prospects for reemployment were low and who were emotionally depressed (7).

The degree of involvement or responsibility of the individuals interacting with the stressor, modifies the biochemical response as shown by Rubin's study of Naval Aviators and Flight Officers during carrier landing qualifications. The pilots who had active flight control showed greater adrenal cortical steroid response and lower anxiety levels than the flight officers whose role was passive(8). Rubin's study also noted that increasing familiarity with the task decreased the biochemical response, a finding also noted for heart rate during simulated landings (9).

Long term memory or experience is a potent modifier of biochemical response as is illustrated by Kramer's study of an 18 hour flight in the F-4C aircraft. Flight duration and difficulty, and flying experience acted jointly to determine physiologic responses to the flight. When difficulty and duration were held constant, flying experience was inversely related to biochemical response (10). With a few notable exceptions most studies of flight stress have been confined to short term situations. Consequently, the long term effects of prolonged stress are only vaguely known. The data that do exist indicate that the effects are far from simple.

The work of Nicholson and others on the effects of aircrew work load has led to several important concepts regarding adaptation to continuous stress. A moderate work load, such as maintenance of continuous operational capability can be sustained in extended missions at the expense of a deterioration in sleep pattern. Such deterioration progressively limits the number of duty hours as the mission is extended and severely limits the crew's ability to achieve optimum performance (11). Nicholson also noted that the high work load associated with difficult landing situations profoundly modified the central nervous activity of the pilot and rendered his assessment of work load more variable (12). This observation raised the possibility that very high work load may lead to central nervous states that impair judgement and performance.

Studies of alterations in schedules and procedures to maintain performance in high work load situations have led to several additional findings. Work load sharing between the pilot and co-pilot reduces the pilot's central nervous excitation but only when the sharing is actuated during the final portion of the let down (12). Work load sharing in extended missions by the use of a double crew does allow for a greater maximum work load but this can be sustained for only a few days and is physiologically very taxing in comparison to staged missions of the same distance (13).

The physiologic cost of extended missions has immediate and delayed components. Fragmentation of sleep patterns during extended missions is a direct result of high work loads and continues into the post mission recovery phase (14). In the post mission recovery phase, there is an endocrine-metabolic depression that lasts 4 to 5 days. The physiologic cost and subjective fatigue experienced depends on the nature and the length of the mission and appears to be relatively independent of the work-rest cycle (13). Noteworthy in several studies of extended missions is the fact that performance does not deteriorate despite increasing fatigue and sleep deterioration (14).

Fighter and attack aircraft rarely participate in extended missions: the more common practice is to fly multiple short duration missions interspersed with periods of nonflying activity. This results in a pattern of very high work load imposed for short periods of time which may continue for days or weeks. Naval aviation entails an additional burden on the pilot in that most of his operational landings will be made on a carrier deck where the margin for error is slight. Inflight electrocardiographic monitoring indicates that regardless of the mission, the highest heart rates occur during carrier landing and secondarily during catapult launch (15) In addition to its obvious survival value, proficiency in carrier landing is a highly regarded skill among naval aviators and is a matter of considerable peer involvement. Each recovery is graded by the Landing Signals Officer and recorded on video tape. The scores for each squadron member are posted in the ready rooms and the televised recovery is reviewed after each mission.

There are few studies of the cumulative psychophysiologic effects of chronic stress and fewer still involving highly skilled occupations such as aviation. There are several clinical studies of psychologic change resulting from chronic combat stress but only a few involve quantitative psychologic and physiologic measures. Field studies of groups under stress have emphasized the use of simple reliable mood measures and biographical data, in conjunction with valid criteria, as being more productive than the use of more complex personality and cognitive measures (16). Even with reliable measures and criteria it is difficult to interpret data from field studies in the light of theories developed from laboratory data. Somewhat arbitrarily, we have assigned the pilot's carrier landing performance as the controlled variable. The cumulative work load, approximated as a function of the product of average flight hours/per man/per day, and the number of consecutive days of flying activity plus the relative danger of the mission, has been designated as the stressor. Biochemical measures and changes in mood and sleep patterns were then designated as controlling variables.

METHOD

Twenty-six pilots and 23 flight officers drawn from two squadrons on an aircraft carrier deployed in the Western Pacific participated in the 6-month study. The aircraft flown by the squadrons was the F-4J Phantom, a two man, high performance fighter. During a nonflying period in transit to the deployment area, the subjects completed a diet and activity form, the Mood Adjective Check List and a biographical form. The subjects also had a blood sample drawn for subsequent determination of cholesterol, and uric, lactic, and pyruvic acids. These data constituted a nonflying control. At a later date, just prior to the first mission flown after a prolonged in-port period, during which little or no flying was done, similar data were collected and constituted the zero cumulative work load level. Similar samples were drawn and data collected after 11 consecutive days of missions flown over nonhostile territory, and again after 22 consecutive days of missions flown over hostile territory where there was significant danger of death or capture. The latter two samples represented respectively moderate and high cumulative work load periods.

Immediately after collection of the blood samples, a 2 ml aliquot of whole blood was precipitated with cold 0.6 N perchloric acid and the sample frozen for subsequent analyses for lactic and pyruvic acid. The remainder of the blood sample was centrifuged and the serum recovered and frozen for subsequent analyses for serum cholesterol and uric acid. Cholesterol was determined by the method of Rubin, et al. (17) and uric acid by the enzymatic method of Liddle (18). Lactic acid was determined by the Rapid Lactate Method (19). Pyruvic acid was determined by enzymatic methods (20).

The Mood Adjective Check List used in this study is a 40 item list that has been factor analyzed and contains six factors: Happiness, Activity, Depression, Fear, Anger, and Fatigue (21). The diet and activity form elicited information regarding dietary content and amount, the use of cigarettes, ethanol and medications and the amount of physical exercise taken in the previous week. The biographical questionnaire elicited age, height, weight, flying experience, specific aircraft experience, carrier experience, education, and information regarding life events and accidents.

The landing performance criterion is that described and developed by Brictson and assigns a score based on the wire number utilized in the carrier recovery (22). Landing performance scores were collected throughout the deployment and the individual scores for the periods under consideration were averaged.

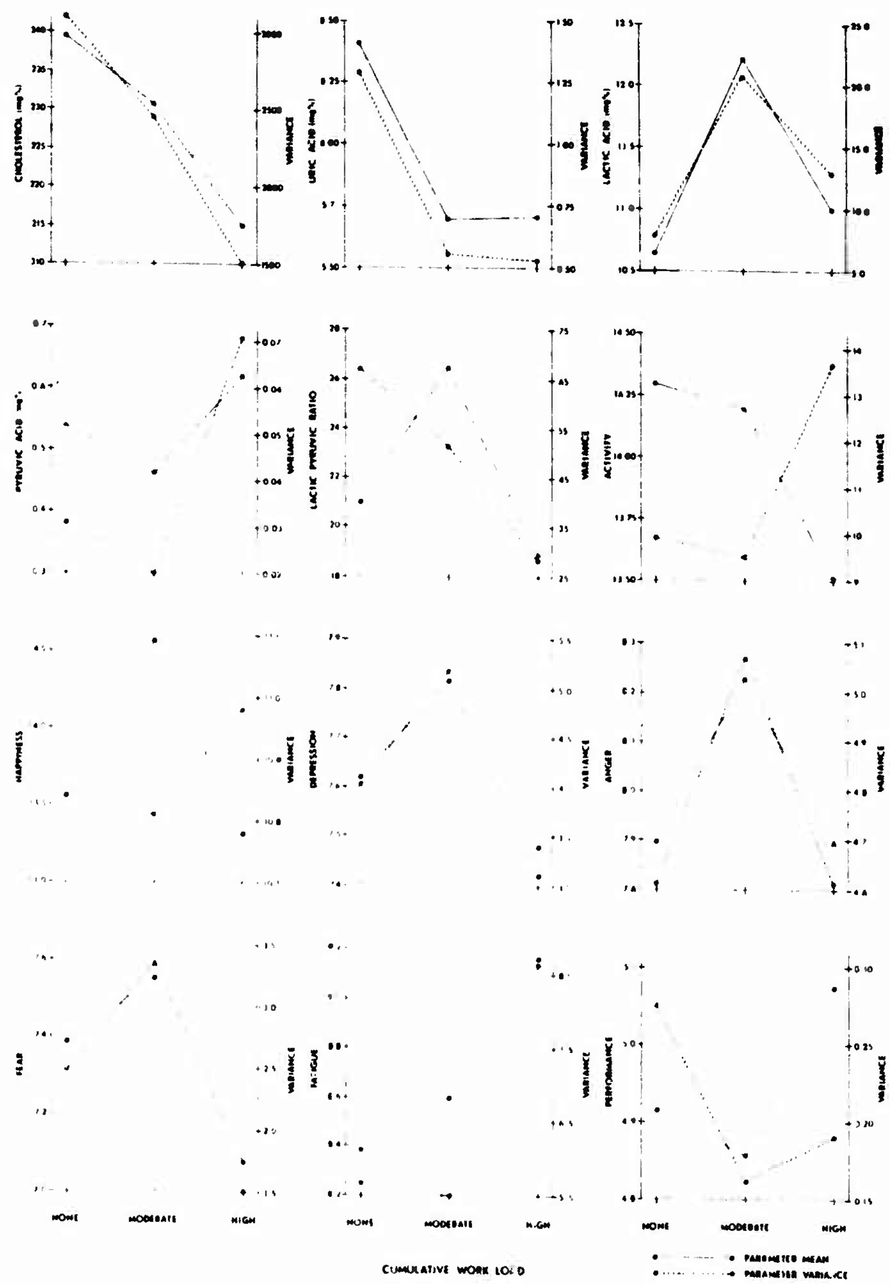
During the period of high cumulative work load, the sleep patterns of a subsample of 27 aviators and a control group of 28 non-aviation related personnel were recorded by sleep logs kept by each subject. The subjects recorded daily the hours of duty, time of going to bed, time of awakening, and responses to difficulty in going to sleep and whether the subjects felt rested upon awakening. From the log data, the total sleep duration per 24 hour period, the duration of sleep episodes, and the time between sleep episodes could be calculated as well as the variability of these measures.

All data collected in the study were transferred to punch cards for computer analysis. Comparison of levels of variables during the periods of the study was by analysis of variance for repeated measures comparison. Differences between variances of the variables at periods of the study was by the F-Max test for uniformity of variance. Relationships between variables were determined by correlations calculated for the different time periods.

RESULTS

The changes in mean levels and variability of the biochemical, mood and performance measures are depicted in Figure I.

Figure I
Biochemical, Emotional, and Performance Changes as a Result of Cumulative Workload



Uric acid decreased significantly under moderate cumulative work load and showed no further decline as cumulative work load increased. Cholesterol level decreased progressively as work load increased and the decrease was statistically significant under the high work load condition. The mean levels of lactic acid, pyruvic acid, and of the lactate-pyruvate ratio did not change significantly under increasing cumulative work load.

The variability of the subjects' serum uric acid levels paralleled the fall in the mean uric acid level and decreased significantly with the onset of moderate work load. The variability remained at a low level under the high cumulative work condition. Similarly, the variability of the subjects' cholesterol levels decreased slightly with moderate work load and fell significantly during the high cumulative work load condition.

Although there was no significant change in the mean blood lactic acid level, the variability of the subjects' lactic acid levels showed a significant increase under the moderate cumulative work load condition. Similarly, despite an unchanged mean blood pyruvic acid level, the variability of the subjects' blood pyruvic acid levels increased significantly under the high cumulative work load condition. Consequently, the variability of the lactate-pyruvate ratio decreased under the high cumulative work load.

The effects of increased work load on sleep patterns is shown in Table IA. The data are taken from a one week period representing the 16th to 18th day of consecutive flying activity and are thus representative of a point between moderate and high cumulative work load. There was no significant difference between aviators and non-aviation personnel in total sleep duration or sleep episode duration. Aviators had a significantly shorter intersleep interval. There was no difference between aviators and non-aviators in the variability of total sleep duration or of sleep episode duration. However, there was far more variation in the intersleep intervals of aviators than of non-aviators. Aviators' sleep was more irregular.

Table IA
Sleep Patterns of Aviators and Non-Aviators Under Moderate Work Load

Variable	Aviators (Mean±S.E.)	Significance	Non-Aviators (Mean±S.E.)
Total Sleep Duration (Hrs/24 Hrs)	7.7±1.5	N.S.	7.46±1.2
Sleep Episode Duration (Hrs)	4.8±1.6	N.S.	5.11±1.8
Inter-Sleep Interval (Hrs)	11.7±3.8	p<.05	14.5±2.6
Sleep Duration Variance	11.67	N.S.	7.61
Sleep Episode Variance	13.28	N.S.	17.13
Inter-Sleep Interval Variance	74.94	p<.05	35.76

Table IB lists the correlations between the sleep measures and the biochemical levels that occurred during the succeeding period of high cumulative work load. Only intersleep interval and variance were related significantly to the biochemical values. Intersleep interval variance was not related significantly to landing performance during the period when the sleep logs were kept, but its relation to performance during the high cumulative work load period has not been determined.

Table IB
Correlations of Aviators' Sleep Parameters with Biochemical Values

Variable	r
Inter-Sleep Interval:	
1. Cholesterol (change from non-flying to High Cumulative Work Load condition)	.417 *
Sleep Episode Variance:	
1. Lactate Level (High Cumulative Work Load condition)	.352
2. Pyruvate Level (High Cumulative Work Load condition)	.357
Inter-Sleep Interval Variance:	
1. Lactate Level (High Cumulative Work Load condition)	.399

*p<.05

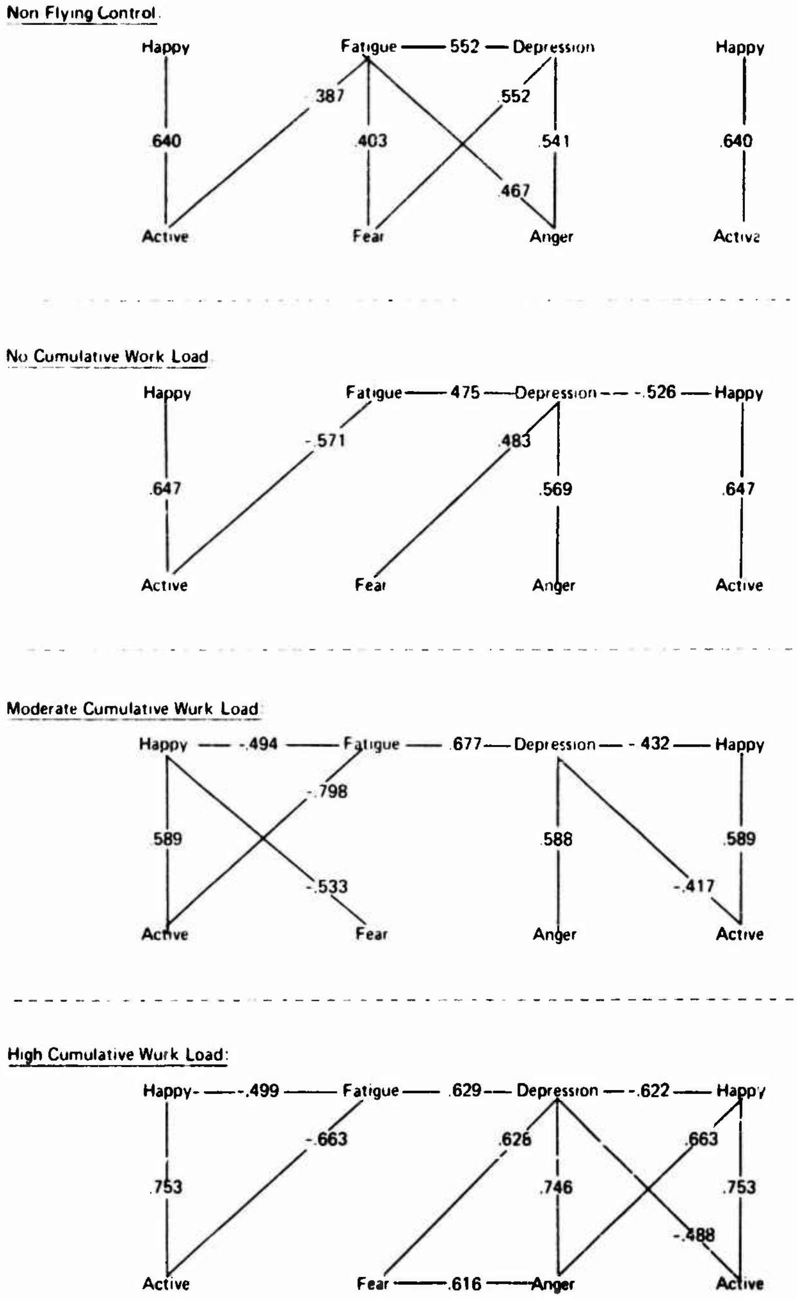
There were no statistically significant changes in any of the emotional states as measured by response on the Mood Adjective Check List. There was a slight upward trend in Fatigue and a slight downward trend in Activity and Fear, as a consequence of increasing cumulative work load. Except for Fatigue, the variability of the subjects' moods did not change as a consequence of increasing work load. The variability of Fatigue response did increase under the high cumulative work load condition, although the increase was significant only in comparison to the nonflying state.

The interrelationship of the various moods is shown in Figure II which also illustrates the changes in mood association pattern as a consequence of high work load. At all times there was a consistent positive correlation between Fatigue and Depression, Happiness and Activity, and Depression and Anger. There was a negative correlation between Activity and Fatigue. Flying activity regardless of work load added a negative correlation between Depression and Happiness to this pattern. Moderate work load gave rise to an inverse relationship between Depression and Activity; this relationship continued during the high work load period. Under high work load the relationship between Happiness and Fear that was present in the moderate work load condition disappeared, and the relationships of Fear and Depression and Anger and Happiness appeared.

The pilots' carrier landing performance scores did not decrease as a result of increasing cumulative work load and in fact, there was a slight, but insignificant upward trend in performance. Also, the pilots' performance was not more varialbe under increasing work load. Indeed, there was an insignificant reduction in variability.

Figure 11

MOOD ASSOCIATIONS AS A FUNCTION OF CUMULATIVE WORK LOAD



——— Indicates statistically significant relationships between moods. The values are correlation coefficients (r).

The relationship of biochemical and mood factors to biographical data was altered by cumulative work load as is illustrated in Table II. Cholesterol which was significantly related to flying experience and age under control and moderate work load conditions, becomes inversely related to height, weight, and accident history under conditions of high work load. Conversely, uric acid under high work load conditions loses the relationship to height and weight that was present under control flying conditions. Under high work load conditions lactic acid becomes strongly related to age, whereas under the same conditions, pyruvic acid loses its relationship to height, weight, and carrier experience. The lactata-pyruvate ratio follows the pattern of pyruvic acid.

There was no common pattern of relationship of moods to biographical data. Happiness, which under non-stress conditions was related to experiential and physical factors, lost all relationship to biographical data under conditions of cumulative work load. Fear under high work load conditions reestablished its initial relationship to specific aircraft experience, and lost the relationship to education that was present under milder conditions. Surprisingly, Fear was not related to accident history at any state. The relationship of Anger to carrier experience was lost under conditions of curulative work load.

Table II
Correlations of Biochemical and Mood Factors with Biographical Data

	Non-flying	No Cumulative Work Load	Moderate Cumulative Work Load	High Cumulative Work Load
Cholesterol:				
F-4J Experience	.145	.275	.290	-.091
Flying Experience	.457*	.498*	.325	.181
ACC-2 Years	.007	-.103	-.201	-.397*
ACC-6 months	-.011	-.015	-.099	-.155
Age	.580**	.710**	.473*	.210
Height	-.172	-.227	-.172	-.577**
Weight	-.130	-.143	-.163	-.406*
Uric Acid:				
Flying Experience	-.190	-.099	-.411*	-.259
ACC-2 years	.058	.367	.261	.060
ACC-6 Months	-.099	.039	.102	.079
Height	.305	.486*	.251	.133
Weight	.309	.455*	.443*	.039
Lactic Acid:				
Age	-.297	-.245	-.273	-.538**
Height	-.198	-.152	.173	.206
Weight	-.134	-.272	.484*	.236
Pyruvic Acid:				
F-4J Experience	.135	.379	-.106	-.228
Flying Experience	-.162	-.071	-.250	-.190
Carrier Experience	.077	.393	-.193	-.158
Age	-.293	-.170	-.155	-.353
Height	.152	-.039	.270	.374
Weight	.064	-.406*	.202	.115
Lactate-Pyruvate Ratio:				
Weight	.140	.428*	.248	.226
Happy:				
F-4J Experience	.476*	.141	.198	.082
Flying Experience	-.306	.377	.341	.121
Carrier Experience	-.386	.135	.114	.098
ACC-6 Months	.481*	.495*	.303	.334
Age	.121	.313	.377	.099
Height	.517**	.225	.269	.048
Weight	.627**	.451*	.313	-.131
Fear:				
Education	.186	-.390	-.397*	-.353
F-4J Experience	-.508**	-.126	-.128	-.422*
Flying Experience	-.349	-.333	-.232	-.346
Height	-.149	-.246	-.227	-.384
Anger:				
Carrier Experience	-.501*	-.560**	-.306	-.249
Height	-.194	-.237	-.346	.225

** p<.01
* p<.05

Table III A indicates the relative strength of the factors determining performance under various work load conditions. On the initial flights, following non-flying duty, experience in the aircraft was the most important factor in determining pilot performance in carrier landing. On subsequent flights as work load accumulated aircraft experience continued in importance, and carrier experience became an important consideration. At high work load, experience was less of a factor and emotional and biochemical influences are seen.

Table IIIA
Correlations of Biochemical, Emotional and Biographical Data with Landing Performance

	No Cumulative Work Load	Moderate Cumulative Work Load	High Cumulative Work Load
F-4J Experience	.536**	.607**	.338
Flying Experience	.311	.166	.378
Carrier Experience	-.162	-.533**	.209
Accident Past 6 Months	-.062	.242	-.067
Accident Past 2 Years	-.192	.038	-.002
Age	-.036	-.115	.225
Cholesterol	-.048	-.089	.374
Depression Mean	-.218	-.081	-.501*
Fear Mean	-.123	.038	-.382

Table III B lists the significant relationships between performance and biochemical parameters at three cumulative work load levels. Several items are worthy of note in this table. The pilots' cholesterol levels were determined by their age and total flying experience. Their uric acid level was a function of height and weight, whereas their lactic acid level was inversely related to age, and was related to the components of the lactate-pyruvate system.

The variability of the pilots' cholesterol values was a function of age, flying experience, and the mean level of cholesterol. However, the effect of age on cholesterol variability was very strong. The partial correlation of cholesterol variability with age indicated that older men have more variable cholesterol levels than could be accounted for by the increased mean cholesterol levels found in older men. The variability of uric acid was closely related to its mean level. The variability of the lactate-pyruvate system was related to the mean levels of the lactate and pyruvate.

Table IIIB
Correlations of Mean Performance and Biochemical Data with Biographical and Emotional Factors Irrespective of Cumulative Work Load

Mean Factor	Correlates	r
Performance Mean	1. Depression Level	-.402*
	2. F-4J Experience	.694**
	3. Total Hours Flown	.412*
Performance Variance	1. F-4J Experience	-.487*
	2. Depression Variance	.360++
	3. Fatigue Variance	.520**
Cholesterol Mean	1. Total Hours Flown	.421*
	2. Age	.546**
Cholesterol Variance	1. Cholesterol	.413*
	2. Total Hours Flown	.437*
	3. Age	.719**
	+4. Age (at constant mean)	.648**
Uric Acid Mean	1. Height	.379++
	2. Weight	.365++
Uric Acid Variance	1. Uric Acid Mean	.505**
	2. Height	.414*
Lactic Acid Mean	1. Pyruvic Mean	.462*
	2. Age	-.526**
Lactic Acid Variance	1. Lactic Acid Mean	.606**
	2. Weight	.493*
	3. Lactate-Pyruvate Ratio	.509**
Pyruvic Acid Mean	1. Lactic Acid Mean	.462*
	2. Lactate-Pyruvate Ratio	-.539**
Pyruvic Acid Variance	1. Pyruvic Acid Mean	.575**

+ Partial correlation
++ Not statistically significant

* $p<.05$
** $p<.01$

Of particular interest was the strong dependence of performance on feelings of depression. Even more striking was the dependence of performance variability on emotional variability, particularly fatigue.

DISCUSSION

These results indicate that the physiologic and emotional cost incurred by pilots of fighter and attack aircraft in combat does not seem to be as severe as that incurred by transport crews flying extended missions (14). Fighter air crews do suffer a deterioration of sleep pattern but this is mild compared to that noted in transport crews on extended missions. There is a metabolic depression manifested mainly as a decline in cholesterol and uric acid levels, which is not great enough to be of pathological significance and may represent a successful, although temporary, adaptation to the imposed stress. An inverse relationship between flying experience and the biochemical response to stress was not present in this study. In contrast to Kramer's study, the stress imposed in this study was comparatively mild, and the effect of experience on performance was strong.

In an analysis of the responses of a group of individuals to a weak stressor, both the mean level of any parameter and the variability of that parameter may be of significance, as both may change as a result of environmental stimuli. Indeed, partition of variance is commonly used as a statistical test of treatment effects. Much of the change in variance associated with environmental stimuli can be attributed to alterations in the level of the parameter affected or of closely related parameters. Thus, the decreases in the variance of uric acid and cholesterol noted as a result of cumulative work load are owing largely to the significant decrease in mean levels of cholesterol and uric acid.

However, changes in variance that are unaccompanied by changes in mean level are a useful and early measure of the response of a group to a stressor, in that they represent the responses of some but not all members of the group. For example, in this study more of the variance in cholesterol is related to age than to cholesterol level and partial correlations indicate that older men have greater cholesterol variability than would be expected simply from the higher cholesterol level that occurs in older persons.

As the variability of the biochemical parameters in this study is related only to the mean level of a particular parameter, or to relatively fixed parameters such as age, or weight, a change in the variability of a parameter in the absence of a change in its mean level is an early indication of the compensatory response of that parameter to a stressor. Thus, the increase in the variability of lactic acid with moderate cumulative work load and the increase in pyruvate variability with high cumulative work load are indications of the pilots' compensatory reactions to the stressor--cumulative work load.

Of the physiological, emotional and behavioral measures considered in this study, cholesterol, uric acid and intersleep interval responded to stress with changes in mean levels and variability. Lactate, pyruvate, and the lactate-pyruvate ratio responded by changes in variability only. Mood measures showed only a slight change in variability, and performance did not change at all in either mean level or variability. On the basis of degree of response to increasing levels of work load, the measures used in this study can be ranked in terms of sensitivity to stress. Uric acid is the most sensitive followed by cholesterol, or sleep measures. Measures of emotion are less sensitive than biochemical or sleep measures but more sensitive than changes in performance.

Even in the absence of significant changes in mean level or variance, shifts in association pattern both within moods and between experiential factors and moods, and between experience and performance, do occur as a result of work load accumulation. Such shifts are not only sensitive indicators of stress response but they greatly complicate the process of performance prediction. Depending on the accumulated flying work load, mood, experience, sleep or biochemistry can enter into the prediction of performance.

Teichman's concept of stress response interposes a compensating mechanism or mechanisms as controlling variables between the stressor and the controlled behavior. From the data in this study it would seem that there is a series of elements in the pathway from stressor to stress response. Correlation coefficients, as a measure of the coordinate variation of two parameters, express the closeness of the relationship between them. From the high correlations between mood factors and performance, it would seem that emotion is one of the main controlling variables of performance. The relationships between lactate and pyruvate and emotion were of low order and insignificant, leading to the hypothesis that some unmeasured variable or variables mediates emotion and the lactate-pyruvate system. The absence of change in lactate and pyruvate levels is to be expected from their relationship to intracellular oxidation-reduction states which would be affected only by a severe stress such as hypoxia (23). Cholesterol changed markedly in reaction to the stressor, but its correlation to lactate and pyruvate was weak, again suggesting an intermediate variable. Of the three sleep measures only one showed response to stress. Sleep measures were correlated with both cholesterol and lactate and pyruvate; sleep may be a parallel to the link between cholesterol and the lactate-pyruvate system.

In summary, Teichman's concept of the organism as a complex self-regulating system, which responds to increasing stress by a sequence of progressive compensatory changes that operate to control behavioral phenomena, seems operationally valid. Cholesterol and uric acid are most sensitive to stress, followed by sleep alterations and changes in the lactate-pyruvate system. Emotion and performance are the least sensitive to stress. In this study we have analyzed several non-contiguous elements of the system. The roles of catecholamine, and adrenal cortical steroids as intermediate variables are currently under investigation. However, the exact role of the various experiential factors is still not clear. Although specific aircraft experience is a strong determinant of performance, other aspects of experience also seem to affect emotional response and biochemical change. The profound modifying effects of experience do not seem to fit Teichman's behavioral model and it may be necessary to develop further theoretical framework to account for the effects of experience.

REFERENCES

1. Teichman, W. H. Interaction of behavioral and physiological stress reactions. Psychol. Rev., 75:271, 1968.

2. Gellhorn, E., and Loofbourrow, G. N. Emotions and Emotional Disorders. New York. Harper & Row, 1963.

3. Thomas, C. and Murphy, E. Further studies on cholesterol levels in the Johns Hopkins medical students: The effect of stress at examinations. J. Chronic Dis., 8:661, 1958.

4. Friedman, M., Rosenman, R.H., and Carroll, V. Changes in the serum cholesterol and blood clotting time in men subjected to cyclic variation of occupational stress. Circulation, 17:852, 1958.

5. Rahe, R. H., Rubin, R. T., Arthur, R. J., and Clark, B. R. Serum uric acid and cholesterol variability: A comprehensive view of underwater demolition team training. JAMA, 206:2875, 1968.

6. Rahe, R. H., Ryman, D., Bennett, L., and Arthur, R. J. Success in stressful training and subjects' serum uric acid and cholesterol. (submitted for publication)

7. Kasl, S. V., Cobb, S., and Brooks, E. W. Changes in serum uric acid and cholesterol levels in men undergoing job loss. JAMA, 205/7:1500, 1968.

8. Miller, R. G., Rubin, R. T., Clark, B. R., Crawford, W. R., and Arthur, R. J. The stress of aircraft carrier landings. Psychosom. Med., 32:581, 1970.

9. Hasbrook, H. A., and Rasmussen, P. G. Pilot heart rate during in-flight simulated instrument approaches in a general aviation aircraft. Aerospace Med., 41:1148, 1970.

10. Kramer, E. F., Hale, H. B. and Williams, E. W. Physiologic effects of an 18-hour flight in F-4C aircraft. U. S. 836th Tactical Hospital, Report No. SAM-TR-66-59. USAF Sch. Aerospace Med.

11. Nicholson, A. N. Aircrew workload. RAF Quarterly, 12:273, 1972.

12. Nicholson, A. N., Hill, L. E., Borland, R. G., and Krzanowski, W. J. Influence of workload on the neurological state of a pilot during approach and landing. Aerospace Med., 44:146, 1973.

13. Hale, H. B., Hartman, B. O., Harris, D. A., Williams, E. W., Miranda, R. E., Hosenfeld, J. M., and Smith, B. N. Physiologic stress during 50-hour double-crew missions in C-141 aircraft. Aerospace Med. 43:293, 1972.

14. Harris, D. A., Pegram, G. W., and Hartman, B. O. Performance and fatigue in experimental double-crew transport missions. Aerospace Med., 42:980, 1971.

15. Roman, J., Older, H., and Jones, W. L. Flight research program VII. Medical monitoring of Navy carrier pilots in combat. Aerospace Med., 38:133, 1967.

16. Helmrich, R. and Radloff, R. Groups Under Stress. New York. Appleton-Century Crofts. 1968.

17. Clark, B. R., Rubin, R. T., and Arthur, R. J. A new micro method for determination of cholesterol in serum. Anal. Biochem., 24:27, 1968.

18. Liddle, L., Seegmiller, J. E., and Laster, L. The enzymatic spectrophotometric method for the determination of uric acid. J. Lab. Clin. Med., 54:903, 1959.

19. Rapid Lactate Fast-Pack. Patented by CALBIOCHEM. 10933 Torrey Pines Rd., La Jolla, Calif. 92037

20. Marbach, E. P. and Weil, M. H. A rapid enzymatic measurement of blood lactate and pyruvate. Clin. Chem., 13:314, 1967.

21. Ryman, D., Biersner, R., LaRocco, J. Mood questionnaire reliabilities and validities (submitted for publication).

22. Brictson, C. A. Analysis of F-4 aircraft day and night carrier approaches. Aerospace Med., 38:12, 1967.

23. Oliva, P. B. Lactic acidosis. Amer. J. Med., 48:209, 1970.

PREDICTION OF PILOT PERFORMANCE: BIOCHEMICAL AND SLEEP-MOOD CORRELATES UNDER HIGH WORKLOAD CONDITIONS

Dr. C. A. Brictson
Dunlap and Associates, Inc.
La Jolla, California

LCDR W. McHugh, MC, USN

and

Dr. P. Naitoh

Navy Medical Neuropsychiatric Research Unit
San Diego, California 92152

SUMMARY

A preliminary longitudinal study of the factors affecting the carrier landing performance of naval aviators under high workload conditions has been carried out. Using stepwise multiple regression techniques, a substantial portion of the variability in landing performance could be accounted for by six factors under zero cumulative workload conditions and by seven factors under moderate cumulative workload conditions. High cumulative workload conditions sharply reduced predictive ability. Although specific aircraft experience and total flight experience were important predictors of average landing performance, blood biochemical levels and emotional states had significant predictive ability. Sleep patterns relate strongly to performance. The factors that determine landing performance change as cumulative workload increases. Suggestions for further research in this area are discussed.

INTRODUCTION

Because of their great theoretical and practical value, studies of the factors affecting pilot performance have interested flight surgeons and aviation psychologists for a number of years. Identification of the skills and factors involved in operating high performance aircraft would facilitate the rational design of successful training programs and would permit the development of selection procedures to assure efficient use of such training. In addition, knowledge of the factors and skills involved in pilot performance could lead to greater safety and efficiency of air operations.

Modern statistical techniques such as factor analysis and stepwise multiple regression are extremely useful not only in the evaluation of relevant factors, but also for the development of predictor equations for experimental testing. The use of such techniques does impose some severe limitations. Stepwise multiple regression requires that a linear relationship exist between the predictors and the criterion variable. In addition, the criterion must be valid and quantitative and the number of subjects must be fairly large (1). Brictson developed a measurement based on the wire number used in aircraft arrestment on a carrier as a measure of landing quality. The use of this criterion in a multiple regression equation identified aircraft handling factors that influenced carrier landings (2). By the use of a similar approach several training variables were found to predict landing quality during night carrier qualifications (3). Other investigators have used multiple regression techniques to evaluate several training variables that critically affect training performance (4).

Several authors have suggested that performance under stressful conditions may be determined by physiologic and emotional factors (5,6). Brictson has noted that carrier landing performance is to some degree determined by the pilot's experience (7). In a previous report, we noted that the relationships between emotional states, biochemical parameters, experience and performance shifted as a result of high workload stress. In that descriptive study, the relative contribution of the various factors to performance was not evaluated (8).

The purpose of this study is to report the use of multiple regression analysis to isolate the factors affecting carrier landing performance under conditions of increasing cumulative workload. In addition, we will describe the relationship between sleep changes and performance.

METHODS

As predictor variables, biochemical, emotional and experiental data were gathered from 26 F-4J pilots deployed on an aircraft carrier in the Western Pacific. As previously described, blood samples for analysis of serum cholesterol, serum uric acid, blood lactate and pyruvate, and assessments of happiness, activity, depression, fear, anger, and fatigue were obtained prior to missions representing zero, moderate and high cumulative workload. Workload levels were determined on the basis of average daily flying hours, number of prior consecutive days of flying activity, and the relative danger involved in the mission. For a one-week period representing a cumulative workload level between moderate and high, the sleep patterns of a subsample of the aviators were recorded on sleep logs kept by the aviators. The logs reported total sleep per 24-hour period, sleep duration, and intersleep interval (8).

The pilots' landing performance scores for the missions flown immediately after blood sampling and emotional assessment were used in the analysis of factors affecting individual mission landing performances. The average landing performance scores for missions flown in the 10-day period following blood sampling and mood assessment were used in the analysis of factors affecting mean landing performance.

The relationship between sleep parameters and individual mission landing performance was examined by correlation. The statistical procedure of stepwise multiple regression was used to identify predictor variables which accounted for significant variance in the criterion. The resultant least squares linear regression equations allowed for the prediction of individual landing performance scores as a function of the predictor variables.

RESULTS

Table IA lists the factors that entered into the predictor equation for individual mission performance during the zero cumulative workload period. The strongest predictor -- history of having had an accident within two years of the study -- accounted for 27% of the variability in landing performance. Addition of serum uric acid and cholesterol to the equation significantly improved the predictive ability. The addition of total flying experience to the equation increased predictive ability significantly but addition of blood lactic acid level did not.

Table IB lists the factors that entered into the predictor equation for individual mission performance during the moderate cumulative workload period. Specific aircraft experience was the strongest predictor of landing performance. The mood factors -- depression and activity -- contributed significantly to prediction. The only biochemical factor to enter -- blood pyruvate level -- increased the predictive ability. Accident history, carrier experience, and anger entered last and raised the multiple R to .927. The seven significant factors accounted for 86% of the variance in mission landing performance.

Table IC lists the factors that entered into the predictor equation for individual mission performance during the high cumulative workload period. The strongest, and, indeed, the only significant predictor was serum cholesterol level. Addition of the serum uric acid level and the lactate-pyruvate ratio to the equation increased the multiple R but provided no significant increase in predictive ability. The three factors accounted for only 42% of the variability in landing performance.

Table IIA lists the factors that entered into the predictor equation for average landing performance during the zero cumulative workload period. Specific aircraft experience entered first and accounted for 28% of the variance in landing performance. The addition of two blood biochemical levels -- blood pyruvate and serum cholesterol -- to specific aircraft experience accounted for 57% of performance variance. The addition of three mood factors -- depression, anger, and happiness -- resulted in a further significant increase in predictive ability. The six factors together accounted for 78% of the variability in average landing performance.

Table IIB lists the factors that entered into the predictor equation for average landing performance during the moderate cumulative workload period. Two experience factors -- specific aircraft experience and carrier experience -- together accounted for about 50% of the variability in average landing performance. Smaller but significant increases in predictive ability were achieved by the successive addition of anger, blood pyruvate level and serum cholesterol level to the equation. The addition of fear and fatigue raised the multiple R to .94 and the seven factors accounted for 89% of the workload conditions.

Table IIC lists the factors that were considered in constructing the predictor equation for average landing performance during the high cumulative workload period. An equation with significant predictive ability for average landing performance could not be developed. No factor could be found that would account for a significant portion of the variability of landing performance. Total flight experience and carrier experience had the highest correlations with landing performance but neither reached statistical significance. The addition of the mood -- depression -- and the addition of the biochemical factors -- blood pyruvate level and serum cholesterol -- raised the multiple R but did not make it statistically significant.

Table III depicts the relationship of sleep pattern and landing performance. Only one sleep measure -- variability of intersleep interval -- correlated with individual landing performance. The correlation coefficient was significant at the .05 level. Thus, the more variable the sleep-wakefulness pattern the lower the landing performance.

DISCUSSION

The relatively small number of subjects in this study limits the power of the stepwise multiple regression procedure. The presence of a small number of subjects can lead to spuriously high multiple correlations, if the data distribution is skewed. The small number of subjects also does not allow the use of a "hold out" sample for cross validation, so that the validity of the predictions so developed cannot be verified. We have used the stepwise multiple regression technique only to tentatively identify those factors that contribute to carrier landing performance under conditions of increasing cumulative workload.

The most striking finding was the effect of high cumulative workload on performance prediction. Under conditions of zero and moderate cumulative workload, more than 70% of the variability in individual landing performance could be accounted for by the predictor variables. Under conditions of high cumulative workload, only 40% of the individual landing performance variability could be accounted for by the predictor variables. The prediction of average landing performance was even more severely disturbed by high cumulative workload. Although our previous report had indicated that there was a shift of the factors associated with performance in the transition from low to high cumulative workload, the absence of significant correlates with average performance during the high cumulative workload period was rather striking (8). Nicholson noted that a pilot's assessment of workload became less predictable under high workload conditions and suggested that either the pilot used a different technique to assess workload under such conditions, or that marked central nervous activity during high workload interfered with the pilot's assessment (9). The data from this study indicate that the situation is more complex. Although a high multiple R could be achieved during zero and moderate cumulative workload periods, the factors contributing to the prediction changed as workload increased. Also, predictive ability abruptly declined in the high workload period. This may mean that, as workload accumulates, different factors determine performance and that the factor contributions may be non-linear.

It is quite clear that one factor that must be considered in performance prediction is workload. Several studies have shown that increases in workload are accompanied by a deterioration of sleep pattern (10) and as these data indicate, irregularity of sleep pattern is closely related to landing performance. At least, irregular sleep is an indicator that the pilots are reaching the maximum tolerable level of sustained performance.

In a previous study we noted that sleep patterns were related to cholesterol, lactate, and pyruvate levels (8). The appearance of those biochemical parameters in equations predicting landing performance may be a consequence of that relationship.

As might have been expected, stable factors such as specific aircraft experience, had greater predictive ability for average performance than did labile factors such as blood biochemical levels. It is not known why experience factors did not appear in equations predicting average performance during the high cumulative workload period. Perhaps the effects of experience on landing performance are inhibited by the deleterious effects of high workload on older, more experienced pilots. In any event there is no firm guide to enable operational commanders to select pilots for high performance, high workload operations.

The small number of subjects in this study and the high multiple correlations achieved emphasizes the need for further research in this area. When the number of subjects is small, the factors that initially enter into a multiple regression equation are likely to be the most valid. These factors should provide the basis for a future study which would consider the effects of sleep, mood, and cumulative workload on landing performance. Such a study should include a large number of subjects so that cross validation of the predictors would be possible.

REFERENCES

1. Darlington, R. B. Multiple regression in psychological research and practice. Psych. Bull. 69:161, 1968.

2. Brictson, C. A., Burger, W. J., & Kennedy, R. S. Predicting the quality of pilot landing performance during night carrier recovery. Aerospace Med. 42:16, 1971.

3. Brictson, C. A., Burger, W. J., & Gallagher, T. Prediction of pilot performance during initial carrier landing qualification. Aerospace Med. 43:483, 1972.

4. Shannon, A. H., & Waag, W. L. Toward the development of a criterion for fleet effectiveness in the F-4 fighter community. Aerospace Med. 44:453, 1973.

5. Frankenhaeuser, M. Behavior and circulating catecholamines. Brain Res. 31:241, 1971.

6. Lazarus, R. S. Psychological Stress and the Coping Process. New York: McGraw-Hill, 1966. Pp 258-318.

7. Brictson, C. A., Ciavarelli, A. P., & Wulfeck, J. W. Operational measures of aircraft landing system performances. Human Factors, 11:281, 1969.

8. McHugh, W. B., Brictson, C. A., & Naitoh, P. Biochemical and mood changes in naval aviators in high workload conditions. Presented to NATO AGARD Aerospace Medical Panel Specialist Meeting, Oslo, Norway, April 22-25, 1974.

9. Krzanowski, W. J., & Nicholson, A. N. Analysis of pilot assessment of workload. Aerospace Med. 43:993, 1972.

10. Hartman, B. O. Field study of transport aircrew workload and rest. Aerospace Med. 42:817, 1971.

TABLE IA

Factors Entering into Predictor Equation for Individual Mission

PERFORMANCE DURING ZERO CUMULATIVE WORKLOAD PERIOD

Factors	R	R^2
Accident History - 2 years	.524**	.274
Accident History - 2 years - Uric Acid Level	.622***	.387
Accident History - 2 years - Uric Acid Level Cholesterol Level	.688***	.474
Accident History - 2 years - Uric Acid Level Cholesterol Level Total Flight Hours	.758***	.572
Accident History - 2 years - Uric Acid Level Cholesterol Level Total Flight Hours Lactic Acid Level	.787***	.619

*** $p < .005$
**$p < .01$

TABLE IB

Factors Entering into Predictor Equation for Individual Mission

PERFORMANCE DURING MODERATE CUMULATIVE WORKLOAD PERIOD

Factors	R	R^2
F-4J Experience	.607***	.368
F-4J Experience Depression	.635***	.404
F-4J Experience Depression Activity	.787***	.620
F-4J Experience Depression Activity Blood Pyruvate Level	.838***	.702
F-4J Experience Depression Activity Blood Pyruvate Level Accident History - 2 years	.865***	.749
F-4J Experience Depression Activity Blood Pyruvate Level Accident History - 2 years Carrier Experience	.892***	.765
F-4J Experience Depression Activity Blood Pyruvate Level Accident History - 2 years Carrier Experience Anger	.927***	.859

*** $p < .005$

TABLE IC

Factors Entering into Predictor Equation for Individual Mission

PERFORMANCE DURING HIGH CUMULATIVE WORKLOAD PERIOD

Factors	R	R^2
Cholesterol Level	.525**	.276
Cholesterol Level Uric Acid Level	.604**	.365
Cholesterol Level Uric Acid Level Lactate-Pyruvate Ratio	.646**	.418

** $p < .01$

TABLE IIA

Factors Entering into Predictor Equation for Average Landing

PERFORMANCE DURING ZERO CUMULATIVE WORKLOAD PERIOD

Factors	R	R^2
F-4J Experience	.536***	.280
F-4J Experience Blood Pyruvate Level	.696***	.481
F-4J Experience Blood Pyruvate Level Serum Cholesterol Level	.753***	.568
F-4J Experience Blood Pyruvate Level Serum Cholesterol Level Depression	.785***	.610
F-4J Experience Blood Pyruvate Level Serum Cholesterol Level Depression Anger	.840***	.700
F-4J Experience Blood Pyruvate Level Serum Cholesterol Level Depression Anger Happiness	.884***	.781

*** $p < .005$

TABLE IIB

Factors Entering into Predictor Equation for Average Landing

PERFORMANCE DURING MODERATE WORKLOAD PERIOD

Factors	R	R^2
F-4J Experience	.607***	.368
F-4J Experience Carrier Experience	.703***	.494
F-4J Experience Carrier Experience Anger	.760***	.578
F-4J Experience Carrier Experience Anger Blood Pyruvate Level	.803***	.652
F-4J Experience Carrier Experience Anger Blood Pyruvate Level Serum Cholesterol Level	.850***	.723
F-4J Experience Carrier Experience Anger Blood Pyruvate Level Serum Cholesterol Level Fear	.884***	.781
F-4J Experience Carrier Experience Anger Blood Pyruvate Level Serum Cholesterol Level Fear Fatigue	.942***	.888

*** $p < .005$

TABLE IIC

Factors Entering into Predictor Equation for Average Landing

PERFORMANCE DURING HIGH WORKLOAD PERIOD

Factors	R	R^2
Total Flight Experience	.378ns	.142
Total Flight Experience Carrier Experience	.439ns	.193
Total Flight Experience Carrier Experience Depression	.497ns	.247
Total Flight Experience Carrier Experience Depression F-4J Experience	.529ns	.280
Total Flight Experience Carrier Experience Depression F-4J Experience Blood Pyruvate Level	.562ns	.316
Total Flight Experience Carrier Experience Depression F-4J Experience Blood Pyruvate Level Serum Cholesterol	.585ns	.342

ns - Not statistically significant.

TABLE III

CORRELATIONS OF SLEEP PATTERNS WITH LANDING PERFORMANCE

Sleep Parameter	R With Performance
1. Total sleep/24 hours	.000
2. Sleep Episode Duration	.000
3. Intersleep Interval	.000
4. Variability of Total - Sleep/24 hours	.000
5. Sleep Episode Variability	.000
6. Intersleep Interval Variability	-.762*

*$p < .05$

GENERAL DISCUSSION ON PAPERS A10 - A13

WILKINSON	I would like to know whether there was a control condition for rising early but not flying?
ULBRECHT	All pilots and subjects had to rise at 0600 hours.
WILKINSON	Did you investigate the partial correlation between sleep interval variability and performance holding experience constant?
McHUGH	No.
WILKINSON	Brictson reported that cholesterol was higher in inexperienced pilots even though performance was raised, but I believe McHugh gave us the impression that the cholesterol might be higher, because it was a symptom of higher effort on the part of the inexperienced pilot to achieve higher levels of performance.
McHUGH	We are not sure what cholesterol represents, but suspect it is a reflection of a metabolic arousal. Experienced pilots tend to have somewhat higher cholesterol levels and so during the period of high cumulative workload the cholesterol values of experienced and inexperienced pilots were approximately the same. We think this means that inexperienced pilots are expending more energy to achieve their performance improvement.
HOLDEN	I was wondering if the regression analysis examined the interactive terms and if any sensitivity studies had been done to establish the stability of the regression line?
McHUGH	We did not examine interactions or stability.
TER BRAAK	Do you consider that to emphasise performance on deck landings avoids important assessments of mission success?
McHUGH	We quite agree that the mission is the important item, but it is difficult to evaluate objectively performance in the overall mission. The carrier landing performance criterion has been validated against a pilots overall skill in carrying out his complete mission.
BRICTSON	We are conscious of this relationship. The assumption of this study was that carrier landing is a critical task and that performance on this critical task at the termination of the mission gives us some idea of what stresses have been experienced. In terms of relating carrier landing performance to other pilot proficiency we are now in fact looking at correlations between pilot landing performance and their combat flying proficiency. Preliminary results indicate that those who are good at landing at night on carriers are also the best in carrying out their mission.
QUESTION	Do you think that sleep logs are an efficient method for estimating quality and quantity of sleep?
McHUGH	For a field study the sleep log is a highly reliable measure which correlates fairly well though not perfectly with EEG recordings of sleep quality.
OBERHOLZ	How did you measure the level of depression and anger?
McHUGH	We measured the level of depression and anger using a 40 item adjective check list. This questionnaire which takes about 30 sec to complete has been validated on 5,000 naval recruits. In addition it has been validated against performance, illness and disciplinary action in naval recruits.
FRANKE	Can you say that a particular landing point on a carrier deck is a normal point and all deviations depend on the pilot behaviour at that moment?
BRICTSON	You have raised the question about random variables entering into any particular landing approach. Except for the conditions which I mentioned yesterday about measures taken under rain etc we have measured these approaches with twin precision radars from 1½ miles down to touch down for over 6,000 landings and found that some variables do enter into the assessment but not in a systematic biasing fashion.
BRUMAGHIM	I was interested in Dr Hartman's presentation of a quantitative index of stress. This is a continuing problem in ground based simulation. Is the School of Aerospace Medicine looking at the use of physiological indices of stress?
HARTMAN	We have biochemical and subjective data which indicates that flying a simulator is equally stressful as flying a transport.
HARTMAN	(in reply to Ulbrecht on correlation between biochemical measures and workload). Sodium and potassium excretion are subject to many factors but they do provide

a useful index and relate to the other measures in an effective way. As far as hormone estimations are concerned we would prefer to measure cortisol. We are also interested in 17-keto steriods.

BRUSHAGHIM Was any recording made of tre illumination level and were tne performance grares broken down for the teletype operator?

HARTMAN We arranged for an environmental survev of the craft before the exercise. The lighting was unsatisfactory. The answer to your second question is no.

FUCHS Did you clarify whether the headache may have been due to a medical condition such as sinusitus?

HARTMAN No, but I think we should have followed this up.

LONG RANGE AIR-TO-AIR REFUELLING
A STUDY OF DUTY AND SLEEP PATTERNS

Wing Commander N.H. Mills RAF
Royal Air Force Strike Command
High Wycombe, Buckinghamshire,
United Kingdom.

Wing Commander A.N. Nicholson OBE RAF
Royal Air Force Institute of Aviation Medicine
Farnborough, Hampshire,
United Kingdom.

SUMMARY

The sleep patterns of ground crew, pilots and tanker crews involved in a long range air-to-air refuelling mission have been related to their duty hours. During such complex operations workload may vary considerably and the demands placed on some aircrew may be very high. It is suggested that the duty hours demanded of individual aircrew should be related to their overall workload. In this way it may be possible to maintain an acceptable sleep pattern in all aircrew and ensure that no individual pilot or crew member is subjected to excessive duty hours.

INTRODUCTION

Long range air-to-air refuelling demands not only sophisticated logistic support but also careful attention to the duty and rest patterns of both air and ground crews. Air-to-air refuelling may be used to transport short range aircraft many thousands of miles within a short period of time and such operations require long and irregular hours of duty both from the receiver pilots and tanker crews. Toward the end of 1969 ten Lightning aircraft were transferred from the United Kingdom to the Far East using the technique of air-to-air refuelling. This paper is concerned with the duty and rest patterns of the aircrew involved in the exercise. Operational aspects have been included if they are relevant to this analysis.

Operational Details. The exercise, which was preceded by about a month of day and night air-to-air refuelling training, was scheduled over two sectors - United Kingdom to the Persian Gulf and the Persian Gulf to the Far East. The aircraft departed from the United Kingdom in formations of 3, 3, 2 and 2 on four successive days and the refuelling aircraft operated from airfields in the United Kingdom, Cyprus, Persian Gulf and Indian Ocean. Supporting ground crews were deployed along the route and pilots not operating the first sector were transported to the Persian Gulf by a Hercules aircraft.

The first sector from the United Kingdom to the Persian Gulf was straightforward and with favourable tail winds the flight times averaged 8.5 hours. The aircraft took off at 0330 hours each morning to ensure arrival in the Persian Gulf before dusk. The first three hours of each sector were operated in darkness. Three refuelling brackets were completed and each formation was east of Malta before daylight broke. There were no obvious difficulties with this plan, but an unsatisfactory sleep pattern preceding the sector and air-to-air refuelling in darkness during the early hours were considered to be significant factors in the induction of crew fatigue.

Two aircraft were diverted to Cyprus, one returning 300 miles from Turkey unaccompanied. Fortunately, the weather was clear and the excellent range of the Andana Tacan beacon made this relatively easy. In general, radio compass bearings in this area were poor and such a recovery could have presented considerable difficulties. The remaining eight aircraft arrived safely in the Persian Gulf but it was clear that, though a flight of about 9 hours was feasible, an instrument approach in poor weather would have imposed a considerable final strain on the pilot. This is why daylight landings were mandatory during this exercise. A thirty-six hour rest period proved essential to refresh the pilots for the next sector.

The departures of all aircraft from the Persian Gulf for the second sector were on schedule at 0530 hours local time. The first hour of the sector was in darkness, but the weather was perfect. About 100 miles northwest of Gan (Indian Ocean) each formation met severe weather conditions and on the worst day a formation of three aircraft logged over five hours cloud flying on this sector. These conditions, together with an instrument recovery in poor weather at the end of an eight hour flight, were found to be a most trying experience for all pilots. Planned fuel recovery figures were affected adversely and in two cases pilots landed from a GCA with limited fuel. Signs of excessive stress and fatigue were obvious at the end of this flight. Six diversions occurred during the sector. Two aircraft were diverted due to a communications misunderstanding concerning the weather and another aircraft because of a serious fuel leakage. The unfavourable weather complicated the refuelling plan specially as far as the three wave formations were concerned.

The unfavourable weather during this sector and the prolonged cloud flying led to pilot disorientation far greater than the pilots believed possible. Most pilots encountered between two and five hours of cloud flying and experienced the 'leans' for periods of between ten and twenty minutes. In some cases the symptoms were severe and increased the sense of fatigue. The tankers appeared to be doing aerobatics, refuelling in the inverted position or flying continuously banked. No loss of control occurred, but scanning of instruments had to be curtailed when in dense cloud to avoid losing the formation. Alterations in heading by the tankers to avoid the worst of the weather did little to alleviate the situation and when flying unaccompanied after the final refuelling several Lightning pilots found themselves in the worst of the weather before any diversion could be made. The operation led to overt fatigue and could have jeopardised safety during the mission. Most pilots realised that they were in no condition to operate for at least twenty-four hours after arrival and that they needed about forty-eight hours to acclimatise to the new environment.

Analysis of Sleep Patterns and Workload. Sleep and duty patterns for all aircrew and some ground crew were obtained from personal diaries. Schedules of five Lightning pilots, six crew members of the tanker aircraft and two ground crew were analysed in detail. Preceding the exercise control sleep patterns, as far as possible, were obtained. The minimum amount of sleep required by an individual pilot over any three day period preceding the exercise was used as an indication of his minimum sleep requirements.

TABLE 1

ESTIMATED SLEEP OF FIVE LIGHTNING PILOTS, SIX TANKER CREWS AND TWO GROUND CREW. UPPER FIGURE UNDER EACH NIGHT IS THE ESTIMATED DURATION OF SLEEP (HOURS) AND THE LOWER FIGURE IS THE AVERAGE OF THE ESTIMATED DURATION OF SLEEP OVER THE THREE PRECEDING NIGHTS. AN ASTERISK INDICATES A DAY OF FLYING DUTY

		Minimum Sleep Required (Hours)	Night of December													
			2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th	15th
Lightning	L1	7.3		4.9	5.8	10.7 7.1	6.7 7.7	2.4* 6.6	3.4 4.2	5.5* 3.7	10.7 6.5	1.9* 6.0	9.8 7.4	11.4* 7.7	9.8 10.3	7.4 9.5
	L2	6.9		7.9	7.9	6.1 7.2	7.7 7.1	2.6 5.4	6.0* 5.4	12.7 7.1	NIL* 6.5	5.3 6.0	7.9 4.4	5.8 6.3	8.5 7.4	
	L3	5.6		8.9	4.6	11.8 8.4	6.5 7.6	7.0 8.4	3.0* 5.5	12.6 7.5	5.9* 7.2	7.0 8.5	5.5 6.1	5.5 6.0		
	L4	5.9		7.5	4.6	9.1 7.1	7.4 7.0	5.5* 7.3	4.9 5.9	7.6* 6.0	7.4 6.6	5.3* 6.0	5.9 6.2			
	L5	7.4		6.1	3.2	NIL 3.1	5.8 3.0	6.4 4.1	4.4* 5.5	4.0 4.9	7.0 5.1	6.7 5.9	6.7 6.8			
Tanker Crew	V1	6.0	4.0*	5.8*	5.8* 5.2	10.6 7.4	6.2 7.5	6.4 7.7	4.3 5.6	7.2 6.0	5.5 5.6	6.5 6.5	5.8* 5.9			
	V2	7.3	6.1*	6.3*	7.4 6.6	8.7 7.5	9.2* 8.4	7.4 8.4	5.5 7.4	8.0 6.9	7.0* 6.8	8.0 7.7	9.3 8.1			
	V3	6.7	7.5	7.5	7.5 7.5	8.0 7.7	5.0 6.8	NIL* 4.3	8.6 4.5	6.7* 5.1	8.9 8.1	2.0* 5.8	3.8 4.9			
	V4	6.6	6.5	3.8*	9.5 6.6	4.6 6.0	5.8 6.6	7.3* 5.9	7.0* 6.7	5.2 6.5	7.0* 6.4	4.4* 5.5	3.5 5.0			
	V5	6.4	7.3*	5.8	7.0* 6.7	7.9 6.9	8.6* 7.8	6.5 8.0	8.7 7.9	8.9 8.0	6.7* 8.1	7.2 7.6	7.0 7.0			
	V6	6.5	4.4*	5.8*	10.1 6.7	8.3 8.1	7.3 8.6	6.4* 7.3	10.3 8.0	8.3 8.3	4.3* 7.6	7.3* 6.6				
Ground Crew	C1	6.5				NIL	13.0	6.4 6.5	8.7 9.4	12.6	8.2 9.8	7.6 9.5	10.4 8.7	8.2 8.7	8.9 9.2	5.9 7.7
	C2	7.3				NIL	6.5	8.6 5.0	8.9 5.0	7.0 8.2	8.6 8.2	7.4 7.7	8.3 8.1	8.7 8.1	8.7 8.6	

During the exercise the aircrew continued to record their duty periods and to estimate the duration of their sleep periods. Total sleep was calculated over three day periods during the exercise to compare with the data derived during the control period. The workload of the aircrew was calculated as cumulative duty hours and related to duration of their individual schedules. The duration of each schedule, for purposes of calculating cumulative duty hours, was from twenty-four hours preceding the first flight to the end of the duty period of the last flight. In the event of a break in flying duty exceeding three days, only the high workload part of the exercise was analysed. The estimated hours of sleep for air and ground crew are given in Table 1. The duty and sleep patterns of the Lightning pilots are illustrated in Fig. 1 - 5 and similar data are provided for the tanker crews in Fig. 6 - 11 and for the two ground crews in Fig. 12 and 13. Aircrew workload during the exercise is given in Table 2.

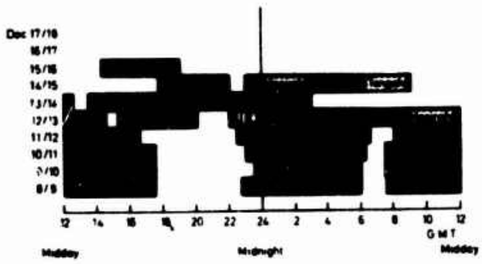


Fig. 1

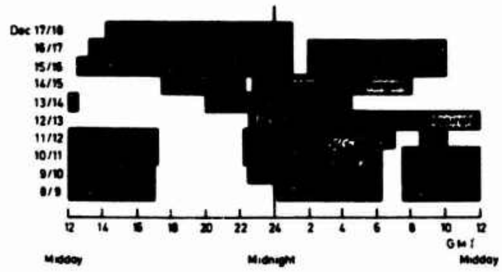


Fig. 2

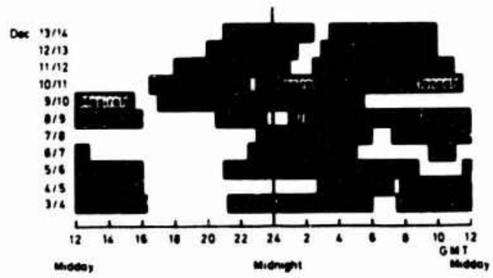


Fig. 3

Fig. 4

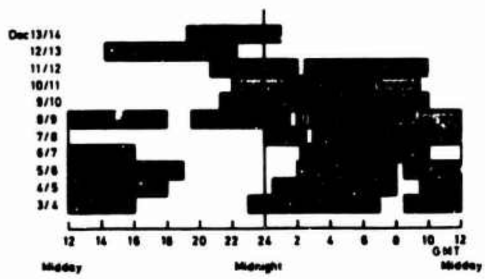


Fig. 5

Figs. 1 - 5. Duty (hatched) and sleep (black) periods of the five Lightning pilots. The horizontal axis covers a twenty-four hour period around midnight. Duty periods involving a flying sector are indicated by departure and arrival airfields.

Fig. 6

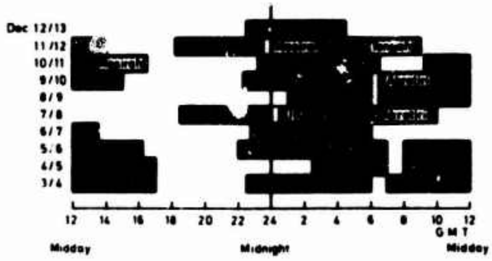


Fig. 7

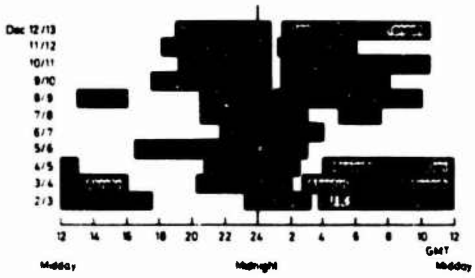


Fig. 8

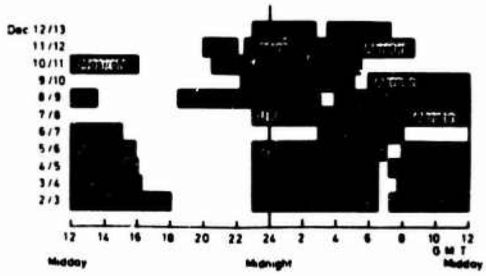


Fig. 9

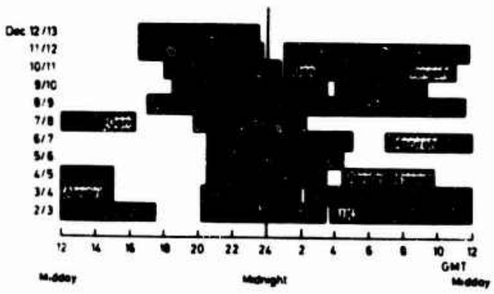


Fig. 10

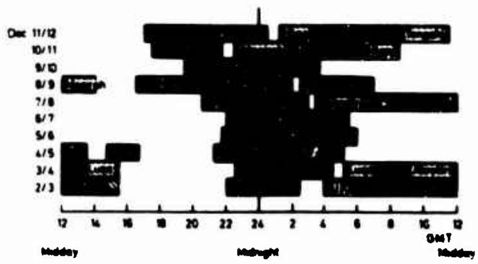


Fig. 11

Figs. 6 - 11. Duty (hatched) and sleep (black) periods of the six tanker aircrew. The horizontal axis covers a twenty-four hour period around midnight. Duty periods involving a flying sector are indicated by departure and arrival airfields.

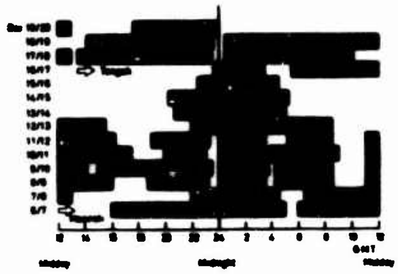


Fig. 12

Fig. 13

Figs. 12 & 13. Duty (hatched) and sleep (black) periods of two ground crew. The horizontal axis covers a twenty-four hour period around midnight. Arrows indicate transit flights.

TABLE 2

WORKLOAD OF AIRCREW

	Total duty hours	Duration of Mission (days)
L1	72.0	6.5
2	36.5	2.3
3	36.0	2.4
4	42.3	4.3
5	52.8	3.5
V1	39.0	2.3
2	42.0	4.6
3	42.5	4.4
4	38.5	4.1
5	46.0	4.6
6	35.0	4.3

DISCUSSION

Severe disturbances of sleep were experienced during this exercise and the unusual duty hours during the days preceding the operation together with the need for an early take-off (0330 hours) led to unsatisfactory sleep preceding the first sector. Not only was the amount of sleep immediately preceding the first sector reduced, but the pattern of sleep during the three day period before the mission was unsatisfactory in some aircrew. One of the Lightning pilots (L5) experienced a thirteen hour flight to the Persian Gulf as a passenger in a Hercules and was required for non-flying duty within sixteen hours of arrival and then operated a Lightning to the Far East two days later.

The workload involved in this exercise (cumulative duty hours related to duration of schedule) was related to previous studies on transport aircrew in which duty patterns compatible with an acceptable sleep pattern had been established. (Nicholson 1970, 1972). An air-to-air refuelling exercise is very demanding and cannot be equated with a normal transport operation, but it is worthwhile comparing the workload of the Lightning aircrew with duty patterns which are known to be only just acceptable for transport aircrew. On the other hand it is considered that the workload of tanker crews can be assessed in the light of experience with transport crews.

It can be seen from Fig. 14 that the calculated workload for the Lightning and tanker crews was close to a workload which was likely to give rise to sleep difficulties in transport aircrew studied previously. In the case of the Lightning pilots the minimum duration of the exercise (from United Kingdom take-off to Far East touch-down) was 2.3 days. Pilots flying the two sectors within 2.3 days were not required to act as a reserve to the other formations. The duration of the other schedules were longer due to additional duty periods or diversions (over six days in one pilot) and the workload tended to be maintained at a high level. Studies on long haul aircrew have shown that workload needs to be reduced as a schedule progresses and this would be particularly important if it was essential to maintain the well-being and operational effectiveness of the aircrew beyond the end of the refuelling exercise.

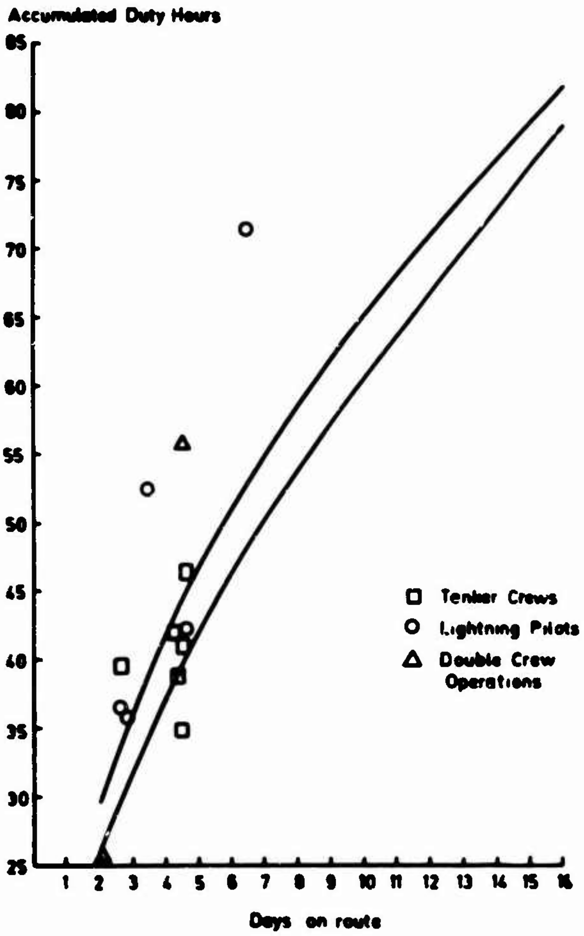


Fig. 14. Workload of Lightning pilots and tanker crews superimposed on a graph relating workload and an acceptable sleep pattern for transport aircrew. The lower line indicates workload which is likely to be compatible with an acceptable sleep pattern. The upper line indicates the maximum workload which may be compatible with an acceptable sleep pattern for transport aircrew. (Nicholson 1972). The workload involved in two double crew continuous flying operations are plotted as triangles (Atkinson, Borland and Nicholson, 1970).

Analysis of the workload of the five Lightning pilots suggested that the schedule would have been incompatible with an acceptable sleep pattern in four of the pilots (L1, L2, L3 and L5) and likely to have created difficulties in the remaining pilot (L4). The sleep patterns obtained from the diaries (Fig. 15) largely confirmed these predictions. Pilots L1, L2 and L5 showed serious disturbances of sleep during the exercise though pilots L3 and L4 maintained satisfactory patterns.

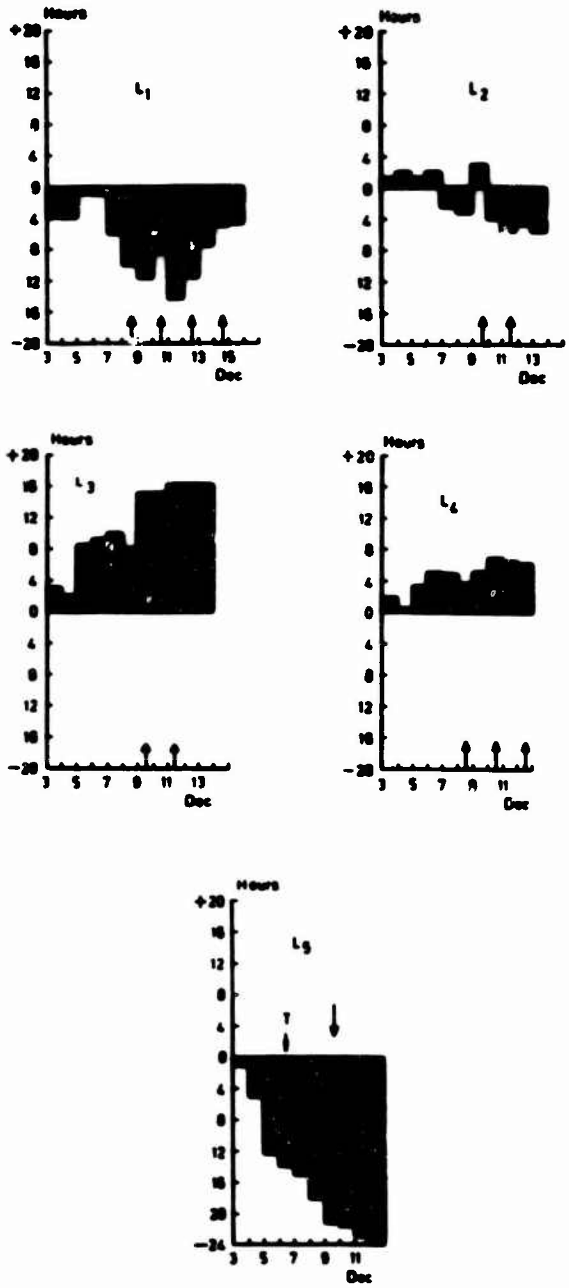


Fig. 15. Estimated sleep gain/loss for the five Lightning pilots during the mission. Arrows indicate flying duty. The arrow labelled T in L5 indicates a transit flight.

The total duration of the schedules of each tanker crew were similar (around ten days), though the workload was concentrated within a four day period. Analysis of the duty periods of the six tanker aircrew suggested that the workload would have been incompatible with an acceptable sleep pattern in two crews (V1 and V5) and that difficulties would have been expected in three other crews (V2, V3 and V4). These predictions were largely borne out by the sleep patterns of the crews (Fig. 16). Satisfactory sleep patterns were observed in V5 and V6, but in the remaining four crews at least some difficulties were experienced.

The transit flight in the Hercules had an adverse effect on sleep patterns (Figs. 12 and 13). One of the ground crew managed to recover an excellent sleep pattern, but the other remained in difficulties throughout the rest of the exercise (Fig. 17).

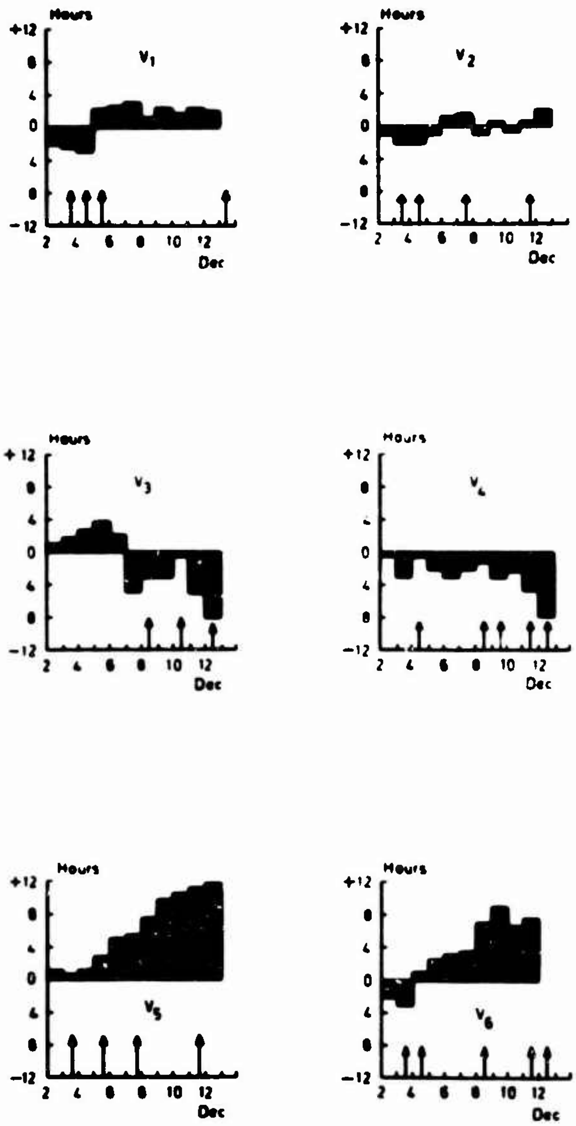


Fig. 16. Estimated sleep gain/loss for the six tanker aircrew during the mission. Arrows indicate flying duty.

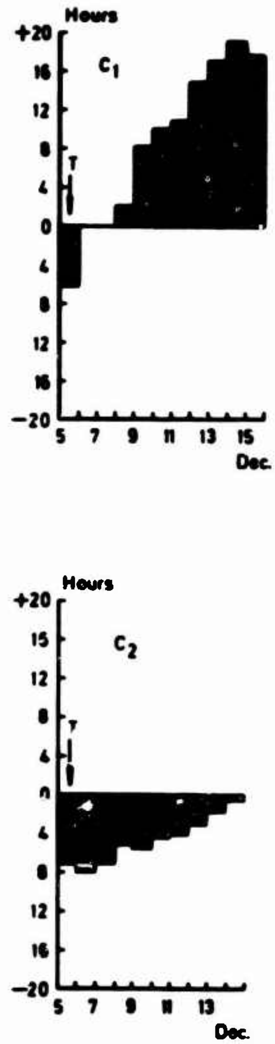


Fig. 17. Estimated sleep gain/loss for the two ground crew during the mission. The arrow labelled T indicates a transit flight.

CONCLUSIONS

This study has explored the relationship between workload and sleep patterns during a long range air-to-air refuelling exercise and has shown that in practice disturbances of sleep pattern could be predicted from a study of anticipated individual workload. Disturbances may affect adversely the well-being and operational effectiveness of aircrew and contribute to fatigue which is likely to be, in any case, a potential factor in this sort of exercise. It is considered, therefore, that careful attention to workload is necessary to ensure that aircrew are not unnecessarily fatigued and that particular attention should be paid to the relationship between cumulative workload and the duration of the individual's schedule.

Studies on long range transport aircrew have shown that unusual patterns of work and rest have an effect on the ability of aircrew to obtain satisfactory sleep and it is suggested that the relationship between cumulative workload and duration of schedules for long range transport aircrew, which is known to be just compatible with an acceptable sleep pattern (Nicholson 1972), could be used, initially, as a model for both tanker crews and fighter pilots. It is appreciated that the workload demanded of Lightning pilots differs greatly from that of long range transport aircrew, but it can hardly be expected that operational effectiveness can be maintained if workload in excess of that which would be tolerated for transport aircrew is demanded.

Studies could be carried out on aircrew involved in particular exercises to determine the nature and significance of unusual duty patterns. Such studies could assist in the planning of operations and may help commanders to predict the overall well-being and operational effectiveness of their crews at any time during a complex operation.

ACKNOWLEDGEMENTS

We are indebted to Group Captain G.P. Black for access to his report on the exercise.

REFERENCES

1. ATKINSON, D.W., BORLAND, R.G. & NICHOLSON, A.N.
Double crew continuous flying operations: a study of aircrew sleep patterns.
Aerospace Medicine 1970, 41, 1121 - 1126.

2. NICHOLSON, A.N.
Sleep patterns of an airline pilot operating world wide East-West routes.
Aerospace Medicine 1970, 41, 626 - 632.

3. NICHOLSON, A.N.
Duty hours and sleep patterns in aircrew operating world wide routes.
Aerospace Medicine 1972, 43, 138.

HIGH WORKLOAD TASKS OF AIRCREW IN THE TACTICAL STRIKE, ATTACK AND RECONNAISSANCE ROLES

by

F. ter Braak, Lieutenant Colonel, RNLAF

SHAPE, Belgium

1. The title of my paper covers the high workload tasks of aircrew performing three distinct tactical roles, namely the strike, the attack and the reconnaissance roles. In the course of my presentation I will cover the total mission, which is from the time the pilot receives his orders until the debriefing following the flight has been completed. At the same time I will also point out the differences between each role, as related to workload, so that a fairly realistic comparison can be made as to the workload of each category.

2. I would like to point out that the task description is based on the pilot flying a single-seater airplane. At the end of my presentation, however, I will include some thoughts on the difference in workload between pilots flying single-seater aircraft and aircrew manning a two-seater, such as the Buccaneer or the F-4.

3. Before going through the missions step by step, allow me to give a general description of each category.

a. Strike. The word "strike" in NATO terminology is usually connected with nuclear missions; in other words, a strike pilot is trained to deliver one of mankind's most lethal weapons. The tactical strike pilot is a lone operator; from take-off onwards, he alone is responsible for the proper execution of his mission. His standard of training and his experience must therefore be of a very high quality. A pilot selected to become a member of a strike squadron will normally have served several years in a fighter squadron executing a different mission. As a secondary task, a strike pilot must also be able to deliver conventional weaponry.

b. Attack. An attack pilot is trained to deliver conventional weapons, such as rockets or bombs on a variety of targets. They range from airfields or radar stations to tank columns or dug-in infantry. The attack pilot normally operates in a formation of four or more aircraft. The proper execution of his mission does not only depend upon his skill in weapon delivery but also on the teamwork attained in the formation of which he forms a part and on the tactics used. He will be trained to perform visual reconnaissance missions and must also be able to execute limited air defense tasks, using guns or cannon.

c. Reconnaissance. The reconnaissance pilot is often described as the eyes of the air and the ground commanders. Flying an aircraft, equipped with several types of cameras, tape recorders and so on, he is tasked to reconnoitre two or three targets and bring back an exact description of his targets, backed up and completed by sets of photographic prints. When making a reconnaissance of army targets, he must give an inflight briefing to the in-flight reporting post, normally collocated with the army commander, so that the required information is available in the shortest possible time. He, like the strike pilot, is a lone operator flying an aircraft normally carrying no defense weapons.

4. Let us now look in detail at the tasks to be performed by each category.

a. Before Flight Period

(1) Intelligence Briefing. Strike: Target briefing and study. Enemy air order of battle. IFF-SIF. Bail-out security. Escape and evasion. Constraints policy. Attack: Target briefing and study. Enemy air order of battle. IFF-SIF. Bail-out security. Escape and evasion. Reconnaissance: Target briefing and study. Required data. Enemy air order of battle. IFF-SIF. Bail-out security. Escape and evasion.

(2) Operations Briefing. Strike: Required flight profile. Recall procedures. Radio procedures. Diversion procedures. Joining instructions. Actions after landing. Attack: Required flight profile. Number of aircraft. Radio procedures. Recall procedures. Type of formation. Tactics enroute. Defense penetration. Target selection. Weapons carried. Tactics in target area. Diversion procedures. Joining instructions. Actions after landing. Reconnaissance: Required flight profile. Recall procedures. Radio procedures. Diversion procedures. Joining instructions. Actions after landing.

(3) GLO Briefing. Strike: Position own troops. In-flight reporting. Own AA and SAM positions. Attack: Position own troops. Target info (for direct support and recce sorties). In-flight reporting. Own AA and SAM defenses. Reconnaissance: Position own troops. Target info (for army targets). In-flight reporting. Own AA and SAM defenses.

(4) Weapons and ECM Briefing. Strike: Weapon settings and checks. ECM procedures. Attack: Weapon settings and checks. ECM procedures. Reconnaissance: Camera settings. ECM procedures.

(5) Flight Planning. Strike: Run-in plot. Map preparation. Extensive weapon release computing. Radar plotting and interpretation. Attack: Map preparation. Radar plotting and interpretation. Reconnaissance: Map preparation. Radar plotting and interpretation.

(6) Time Available. Strike: ± 3 hours. Attack: ± 1 hour. Reconnaissance: ± 1 hour.

b. During Flight Period.

(1) Base to Target Area. Strike: Take-off and leaving procedures. Map reading. Radar interpretation. Flight plan adjustments. ECM checks. Instrument checks. Look-out and radar scan. Attack: Take-off and leaving procedures. Map reading. Radar interpretation. Flight plan adjustments. ECM checks. Instrument checks. Formation keeping. Look-out and radar scan. Reconnaissance: Take-off and leaving procedures. Map reading. Radar interpretation. Flight plan adjustments. ECM checks. Instrument checks. Look-out and radar scan.

(2) Target Area. Strike: Run in. Weapon settings. Highly accurate navigation. Execute attack maneuver. Highly accurate flying. Escape maneuver. Attack: Run in. Weapon settings. Highly accurate navigation. Pull-up. Target recognition and selection. Attack. Escape maneuver. Reconnaissance: Run in. Camera settings. Highly accurate navigation.

(3) Return Flight. Strike: Map reading. In-flight reporting. Diversion procedures. Joining and landing. Attack: Map reading. Visual recce sortie. In-flight reporting. Diversion procedures. Joining and landing. Reconnaissance: Map reading. In-flight reporting. Diversion procedures. Joining and landing.

(4) Time Available. Strike: ± 1 hour 20 minutes. Attack: ± 1 hour 20 minutes. Reconnaissance: ± 1 hour 20 minutes.

c. After Flight Period.

(1) Debriefing. Strike: Intelligence. GLO. OPS. Attack: Intelligence. GLO. OPS. Reconnaissance: Intelligence. GLO. OPS.

(2) Time Available. Strike: ± 30 minutes. Attack: ± 30 minutes. Reconnaissance: ± 40 minutes.

d. Total Time Available. Strike: ± 5 hours. Attack: ± 3 hours. Reconnaissance: ± 3 hours.

5. Now let us consider how to translate the previous tasks descriptions into workload. Total workload of a tactical pilot is built up out of three factors, the physical, the mental and last, but by no means least, the psychological one.

a. Physical. At first glance the direct physical factor does not seem overly large. Flight preparation, consisting of briefings and flight planning, takes from approximately 1 hour for attack and reconnaissance pilots to 3 hours for a strike pilot. This should pose no problems for a fit man. A flight time of say 1½ hours, the average duration of a tactical mission, may also not impress for instance a transport pilot used to flying trips of 6 hours or more. In the case of a tactical pilot, however, several subfactors lead to a high increase of the physical factor. In many cases a tactical mission is carried out at a very low level. Altitudes between 100 and 500 feet over various types of terrain with speeds ranging from 400 to 600 knots are normal. The terrain avoidance action will constantly have to be taken at these altitudes and speeds, serious buffeting will often be experienced resulting in G-forces from minus two to plus two. Maneuvering in the target area may require G-forces of up to 6-G. To maintain the required formation keeps the attack pilot normally fully occupied. Added to this the physical effort required to perform defensive maneuvers in case of fighters, it must be clear that the physical factor in toto is not as small as at first glance it seemed to be.

b. Mental. The mental effort required to perform a tactical mission forms a large part of the total workload. The briefings and flight preparation, as described in paragraph 4, require from 1 to 3 hours of highly concentrated work. During flight the mental effort required to perform the low-level navigation and to readjust prepared flight plans because of unsuspected weather deterioration, technical problems, and/or enemy action will demand a very high mental capability.

c. Psychological. The psychological factor as it relates to the total workload is of course difficult to measure as it varies from person to person and also from day to day. The influence this factor has on the total workload depends on many things such as standard of training and total experience, physical condition, personal problems and many others. Whatever the situation may be the psychological factor has a direct influence on the output the pilot or aircrew member can produce and therefore on the total workload.

6. Training. In the past minutes I described the multitude of tasks the tactical pilot is required to perform in the course of his mission. From the total time available to for instance an attack pilot, it must be clear that many of the actions described have become more of a routine matter thanks to the training received. The more intensive the training and consequently the higher the proficiency standard, the more influence this has on the total workload. Whether this training is executed by the actual execution of simulated wartime missions or by simulation is a matter which will be discussed at a later stage of this meeting and I will therefore refrain from discussing this at the present.

7. Two-man Crew. As stated at the beginning of my presentation, I would now like to offer a few thoughts on workload for two-man crew. At a first glance, one could draw the conclusion that as two qualified men are available to carry out the tasks described the workload per crewman if not halved is most certainly reduced by a fairly high factor. This however is far from the truth. The physical factor is not reduced as each crewman performs his task under the same circumstances as a single pilot. Usually, however, this factor is somewhat higher because the missions accomplished by a two-seater crew are often more complex requiring greater physical effort. The same argument covers the mental factor. Thirdly, the need for crew coordination during the mission is a task in itself keeping workload high. And fourth, the psychological factor which is equivalent in both cases. In closing, I believe I am right in saying that in general we cannot speak of an appreciable difference in workload between the two categories.

DISCUSSION

POLLARD

I would like to add to the presentation a couple of more stresses which we have found to be particularly important. The first one is that of extended missions of up to 6 hours which involve 2-3 hours of air-to-air refuelling and the second is the difficulty of communications over tactical crowded frequencies.

TER BRAAK

Yes. I quite agree with you but I was unable to deal with these particular problems in my presentation.

NICHOLSON

In dealing with the problem of single and double crew aircraft I would like to link this to the question which you raised about creating an optimum performance under the very high workload of the acquisition and delivery phase. Is it possible that a two man system could provide a better performance during the very high workload situation of the acquisition and delivery?

TER BRAAK

I think that a two seater aircraft would definitely be an improvement in the total performance. The workload is not lower for each man, but the performance could well be improved. The more complex missions nowadays take a lot out of you. You cannot fly more than one mission a day. Possibly with a two man crew we could increase the number of missions to one and a half a day.

FRANKE

From the standpoint of research and looking to the cost effectiveness of an air force I have no indication that a two seater aircraft is more cost effective than a single seater. I would like to question our U.S. colleagues. Do they have information that a two seater aircraft is more effective?

POLLARD

It seems to me that the A7 is just as effective as the F4 if it is an attack mission, but in many instances there is equipment aboard the F4 that the A7 could not carry, such as laser designators. So there are missions which the A7 is not equipped to fly. It depends on the mission and the equipment aboard the aircraft. I would agree that with a two seater aircraft you could perhaps extend the mission and the roles of the aircraft.

THE AIR DEFENCE ROLE

Wing Commander J. Hutchinson RAF
Royal Air Force Binbrook
Lincoln
United Kingdom

Introduction

In considering the workload involved in the air defence role, it is important to be clear as to the dimensions of the subject. On the one hand, any air defence situation will pose a problem which can be defined in absolute mathematical terms, and whose solution implies a certain degree of effort from the defending pilot or crew using given equipment. However, the essence of air defence is that the intruder will usually set about making the defender's task as difficult as possible, the tactics he adopts, in addition to making the problem less tractable, may also reduce, through stress, fatigue, fear or several more factors, acting singly or in concert, the ability of the air defence crew to solve the problem. The perceived workload facing the crew whose ability is thus impaired may prove to be beyond their capability, even though in absolute terms the problem is not overwhelming.

My specialisation is the air defence of the United Kingdom, and any examples I cite will be based upon operations conducted in that role. I must also be careful to define 'Air defence' as the finding, and shooting down of airborne enemy intruders using manned interceptors. Some would claim that the air defence role includes all means used to reduce an enemy's offensive air capability, but we are concerned here solely with the difficulties experienced by defending air crews.

In the post-nuclear-tripwire age, air defence covers an enormous span of possibilities, against all of which the defending crew has to maintain a capability, because it is the enemy who decides the tactics. In the Royal Air Force we train our pilots to operate round the clock in defence of the United Kingdom; in defence of Mediterranean and Central European bases, and in defence of the fleet at sea; and for all these roles we train against all possible threats from supersonic air-launched missiles at high level, through high subsonic bomber penetrations at low level, to air combat manoeuvring against fighter-type aircraft. This range of possibilities includes head-on, beam and rear attacks with missiles, guns or both, pressed home either semi-automatically using electronically-computed steering information, or visually.

Definition of Workload

The total load on the crew operating in this role is the sum of several different pressures - physical, mental and emotional being the most obvious. Unfortunately, in air defence the enemy by definition holds the initiative, and can build up pressure on the crew in any or all of these areas, conceivably without firing a shot. Physical or physiological pressure is a product of the crew's working environment and equipment. Air defence aircraft are normally designed to achieve a performance advantage, and a price has to be paid for this. Rates of change of height achieved are very large, 50 - 60,000 feet per minute not being uncommon for short periods. Most modern fighters have the thrust, at any rate at low level, to sustain very high g loadings indefinitely, and they can usually approach 1G linear acceleration in level flight. Altitude changes are large, frequent and very fast, and, particularly when combined with rapid speed changes, they place considerable strain on the crew's sensory mechanisms.

Also in pursuit of performance, cockpits can be very small, resulting in cramped conditions, with instrument presentations miniaturized to the point where interpretation requires much effort, if they can be seen at all. The view from the driving seat can also be severely restricted, making for more physical effort to maintain good lookout. And lastly, I have never flown an aircraft in which it was always possible to achieve a comfortable temperature throughout without excess noise and commotion. All this adds up to an appreciable physical strain on the crew, before a single shot has been fired, and it undoubtedly makes for a high degree of fatigue.

Then there is the pressure of the emotions. Fatigue is the commonest, and most enervating, but every air defence pilot knows the debilitating effects of hunger, pain (usually from the bladder, or the seat, but occasionally from cold feet, hot feet, cold hands or jarred elbows) and apprehension or even fear. Disorientation, though physiological in origin, is a very effective trigger for the main and standby adrenalin pumps and is most common on dark nights at low level and high speed. The art of mastering these forces is a part of the training of every air defence crew, for the better they can overcome them, the more of their total intelligence they can devote to the solution of the interception problems they face - the mental pressure.

The Mental Pressure

Air defence as we are considering it consists of seeking out the enemy in the air and destroying him. Before this end can be achieved, a good many links in a delicate chain have to be completed. First, the defenders have to become aware of the enemy's presence and divine his intentions. Then the defending pilot has to get airborne himself; next he has to navigate to the vicinity of the enemy and make contact with him using his own sensors. Then he has to manoeuvre into a firing position, and finally, he has to select a weapon and discharge it from a position from which it will destroy the enemy. If I tell you that, for the pilot defending the UK, the whole of that sequence can take place 50 miles from his base, at supersonic speed in the stratosphere, within 15 minutes of his being fast asleep in bed you will have some idea of the problem he faces.

In theory, the problem yields to mathematical solutions. The ground station picks up the enemy and a computer calculates his course and speed and scrambles a waiting fighter towards the computed point of interception. Knowing the fighter's performance, the computer can tell the pilot exactly how to navigate until he picks up the intruder on his own sensors. Having locked his radar to the intruder, the fighter pilot can be presented with steering information to manoeuvre him into a firing position, computed from the accurately measured angles and ranges supplied by his own radar. All he has to do is to launch a weapon when the computer tells him to.

However, all this theory collapses before three unfortunate facts. Firstly the enemy may successfully avoid early detection by friendly sensors; secondly having been detected he may take steps to confuse the sensors which are tracking him, and thirdly he may seek by manoeuvring to complicate the defender's problem in attempting to shoot him down. In attempting to solve an interception equation in which every one of the variables is changing rapidly, the pilot's mental capacity, which is the only alternative to mathematical computation which allows qualitative judgements, is taxed to the full. For one thing, by delaying his discovery until the latest possible moment, the enemy compresses the resultant stalk and interception into minimum time, speeding up the rate at which the defender must work; and for another, by producing false electromagnetic waves to compete with the genuine ones initiated by the defender. He can cause the latter to have to base some finely balanced decisions upon a wholly subjective weighting of the perceived evidence.

Consider the components in the total skill of the air defence crew. Firstly, they have to get airborne. If warning is available early enough, they will be well briefed before they go, usually by secure means, that is their firm baseline. If warning is denied or may be denied, they will go on an estimate, to sit astride the enemy's most likely line of approach. In either case, they must be prepared for just about anything, because they can only react to events. Once airborne, they are entirely dependent on the propagation of electromagnetic waves for further instructions and for early detection (the only alternative, the human eyeball, is very restricted in range indeed). They carry several sensors, both active and passive, and if they hope to achieve surprise they will initially rely only on the passive ones. Unfortunately, none of the passive sensors is able to distinguish friend from foe, and the enemy, if he judges it tactically sound, will try to fill them with false information, spurious and incorrect radio instructions being one obvious example. If the crew has to switch to active sensors, the enemy will be made aware of its presence and any attempt to lock on will signify to the enemy that he has been detected clearing him to transmit his full range of electronic deception measures.

On the way to the target area the crew will navigate either by inertial means or on instructions from the ground controller. If neither of these is available the only means is dead reckoning, aided by aircraft radar when pointing at, and within range of the coast, which in the case of the UK is away from the enemy. From now on, dead reckoning navigation has to be elevated to the level of a precise science. An error of 20 miles built up over an hour, which is about 5% of distance travelled, could make the difference between getting home on minimum normal fuel and running out - and in the UK all fighter bases are close to the coast, so there are no drop-short alternatives.

Once on station, the crew of a UK air defence aircraft has the further navigation problem of in-flight refuelling. Many air defence missions rely on this technique, which requires first that they find the tanker. Here the problem is three-dimensional; it is necessary to leave station with enough fuel to reach the tanker's assumed position, and to arrive there with enough to go home if the tanker is not there or unable to supply. Most air defence refuelling missions are conducted in radio silence, so the crew can draw only on its own cunning and experience to find the tanker either after an interception or from a patrol line. There is one further complicating factor in this equation; fuel consumption rates vary widely according to altitude and speed, being roughly doubled for supersonic flight. Whilst a computer can be of help in solving many of the attendant problems, a properly balanced solution can only be devised and kept up to date by the crews acting as much on intuition as on any scientific process.

Apart from navigation, the crew has to wrestle with the problem of detection. Here there are two difficulties; the ground controller can usually see what is going on, but he has to transmit it to the crew by electromagnetic means which the enemy will try to hinder; and the crew has to manipulate its own airborne sensors to pick up the enemy, who will do his utmost firstly to avoid such detection and secondly, once detected, to reduce the quality of information that the defender can glean about his movements.

Until somebody invents a foolproof unjammable radio, communication between controller and crew is bound to be difficult. It can not be entirely prevented but it requires continuous frequency changes and the interpretation of other information such as intercepted jamming signals, and takes both time and mental effort and the resultant information will certainly be of reduced quality.

Radar operation and the countermeasures to which it may be exposed is a subject on its own. Essentially, a computed attack requires continuous range and angle information on the enemy aircraft. If either of these is denied - and a whole range of means is available, each of which has its own counter-play-the computation is transferred to the pilot's head. When you are considering attacking a target on which you are closing head on at 1000 knots, an error of 5 seconds in timing can mean an error of a mile or more in firing range; clearly in spite of other distractions, a high level of ability both to read a very small radar screen, and to make complicated calculations quickly in the head, is essential. The enemy, of course, will do his utmost to make the radar screen unreadable and only a human eye can separate the genuine from the spurious.

Next, the problem of weapon release. Modern missiles can be very effective, provided they are fired from within what is called a success zone - a pocket of sky of definable volume within a certain distance of the quarry. Unfortunately, the enemy if aware of the fighter's presence, will certainly take steps both to prevent him from computing the point of entry to the success zone, and to make it as small as possible. The success zone can be computed from the maximum and minimum ranges of the missile, the height of launch, the height and angular differences between fighter and target, the relative speeds, rates of turn, climb and descent and the navigational and fusing capabilities of the missile. The denial or corruption of any one of these parameters, before launch, or their subsequent alteration, can result in a falsely computed launch point, and of course a resolute enemy will take steps to ensure that he achieves just that, both by use of ECM and by evasive manoeuvre. The crew's task is therefore to be ready at any time to compute the weapon release point mentally, a task for which the time available is usually measured in seconds, and which has to be conducted not in an armchair beside the fire, but in a dynamic situation with superimposed distraction supplied from a variety of sources, not the least of which may be the target's own defensive armament.

After all that, the crew has only to find its way back home, or on to its station, or to another task, which may seem childishly simple and probebly would be if only its brains would work, the sun would rise and the weather was fine.

Operational Equipment

I should now like to illustrate what I have been saying by relating it to the equipment we use. 90% of the RAF's air defence force is provided by various marks of Lightning all of which are externally fairly similar but which differ in fuel capacity, instrumentation, equipment, weapons and fire control systems. The Lightning was conceived 25 years ago at a time when the relatively low-thurst engines available were the limiting factor in achieving the desired Mach 2, and as a result the unique layout was dictated by the need to keep the frontal area as small as possible. The resultant aircraft has a very good performance, even today and flies like a real aeroplane even though it doesn't look much like one. But the internal stowage volume is minimal and the cockpit very cramped even for the single occupant, even the view out being restricted by the design of the canopy. The flight instrumentation is unique, a miniaturized integrated display which combines into five dials the information normally spread over ten. There is a good autopilot, or flight control system, whose aid is very necessary when the pilot has to concentrate on other tasks, for like all good fighters the aircraft is of low aerodynamic stability and is sensitive to small changes in pitch angles. The aircraft will accelerate at about 10 knots a second all the way to sonic speed at sea level, and will reach the tropopause from rest in about two minutes. All marks have AAR facilities and can stay airborne for up to 11 hours.

The aircraft's navigation facilities are rudimentary but reasonably effective under favourable circumstances. It has TACAN with an offset computer, that is the ability to 'move' the ground beecon about by electronic means to put it anywhere it is required. TACAN is of course, susceptible to corruption by jamming. The aircraft's radar has a coast-mapping facility out to 60 miles, or 140 miles using second time-base returns, which is useful if it happens to be pointing at a coast but not otherwise. Apart from that, navigation has to be conducted by dead reckoning, or on instructions from the ground, except that it is possible to home onto a radio transmission by using an anti-jamming device, a facility we sometimes use to find both jammers and tankers.

The Lightning's integrated radar and fire control system is based on a simple and reliable pulse radar augmented by computers which generate electronically-displayed instructions. The radar aerial can scan an area of sky 80 miles wide by 20,000 feet high at 40 miles range once every two-thirds of a second, and it has a wide range of selective manual control facilities. Control is under the left hand, on a single handle containing 14 different controls which perform 18 separate functions. Interpretation of this display is by eye a process which is analogous to reading a book the learning of it being one of the earliest lessons undertaken by the aspiring fighter pilot. The screen is only about 4 inches square so precise interpretation is an art which takes time to acquire; we always practice the worst case, which is that of a return appearing on the tube without any supporting information being supplied by the ground controller. In that case, the pilot has to assess its relative height, speed and heading by a process of mental arithmetic, and act accordingly, basing his actions on an intimate knowledge of his own aircraft's performance. When I tell you that a supersonic rear interception can take as little as 45 seconds from initial contact, you get some idea of the rate at which the pilot has to work. In another case, a high flier interception started from 30,000 feet below the target, can take as little as two minutes. In the extreme case, a head-on interception against a supersonic target can take 20 seconds or less from initial contact to weapon release, but in this case locking on is essential. Even so, in that time the pilot has to spot the target, highlight it manually, compute and make an initial course correction, lock on (an action requiring four separate actions performed with the left hand itself and three of the fingers), follow the steering instructions, monitor the locked-on indicators, compute a breakaway manoeuvre, check the missile acquisition and programming and press the trigger.

One other radar function is the visual identification of unlit targets at night. This exercise which may require closure to a few yards range on an evading target at low level, requires extreme delicacy of aircraft control in response to information displayed on the radar tube, and of course, can not be automated in any way.

When it comes to firing his weapons the pilot has a choice between locking on and giving his computers the task of calculating the missile success zone, or of sighting visually and doing the computation himself. In the first case, provided the enemy has no means of breaking or deceiving his lock he need only monitor the computer's solution and execute the steering instructions if they appear to be reasonable. In the second, he must compute the solution himself using his judgement and whatever facts he can acquire; but of course he makes his adversary's ECM task very much more difficult.

Whilst all this is going on, the pilot of the Lightning has also to keep abreast of his fuel state, the tactical situation, and the state of his home base and diversion weather. He also has to operate his secondary radar transponder to changing codes to ensure that ground stations can identify him in the face of any enemy attempts to compromise the system.

Conclusion

I have said little enough, but I hope I have given you some idea of what is involved in the air defence mission. War of course, has always been a contest between people not machines, and that remains true today. The problem of today is that the dynamic situation has speeded up so much that the load on the individual has reached proportions which threaten to engulf him. One answer to that in air defence is the two-man crew, a solution which we embraced years ago but abandoned temporarily with the Lightning in response to the seductive wooing of those who believed that all things can be solved electronically. With the introduction of the Phantom in place of the Lightning in the UK air defence force, another brain and set of senses will be available in any situation; but one thing remains certain, and that is that in the heat of the chase the crew will require all its faculties in full working order if it is to keep abreast of the problem.

DISCUSSION

JEX

Would you comment on the possibility of remotely piloted interceptors for the Air Defence role?

HUTCHINSON

I believe that current research is concerned more with the air combat situation than with the bomber interception mission which forms our primary task. I would say that unless you can find some means of propagating secure electromagnetic signals for a remotely piloted vehicle there are very considerable difficulties. There has always been a counterploy for every ploy you put into effect and this is certainly true today. If we are on the verge of a break through in completely secure signal propagation I would be enthusiastic, though I would hate to think of operating a fighter from behind a desk underground somewhere. I wouldn't quite get the feel for it!

HOLDEN

I imagine you go through a number of simulated missions in the year with real flights. Do you assess the performance of these missions and what are the prime causes for mission failure?

HUTCHINSON

We certainly run through full scale missions both in ground simulators and in the air. The major recording facility we use is a camera which records the pilot's view of the radar scope. It does not record any other parameters. Failures come about as a result of the breakdown in any one of the links I mentioned in what is a delicate enough chain. I wouldn't like to put my finger on any specific one which is any more common than any other, except to say that the threat of the low level attacker remains by far the most difficult with which we have to cope. This is the area, naturally enough, in which we get most failures and they are usually through inability to detect in sufficient time. A low level attacker doing high subsonic speed is a pretty tough nut to crack in the interception business. If you turn in behind, outside firing range, you are committed to a long tail chase and that is something we wish to avoid. A tail chase of more than 40 to 50 miles leaves a hole in your defences.

HOLDEN

Can you give a clear identification of pilot failure as causes for borderline or failed missions?

HUTCHINSON

Pilot failure usually comes about as a result of degradation of ability. This is usually cumulative as a result of a number of factors of which I would say the most common is that the dynamic situation will simply prove too fast. For example, in the case of a low level interception at night with minimum warning time failure will most always result. This is a combination of the pilot's inability to pick-up the intruder soon enough and the difficulty of the control agency (airborne aircraft or ship) in making contact with him and giving to him the information he requires.

HOLDEN

Are most of those factors dependent on the availability of the information within an adequate period of time and not on the ability of the pilot to operate on the basis of that information?

HUTCHINSON

That is perfectly true. I think that all nations are suffering from a wide gap in their low level early warning facility. This is the ability to get airborne and to do what you have to do. I should add that in the Lightning the performance of the aircraft sets its own problems at times, because, as far as I know, it is the only aeroplane currently in service in the Western world in which you can attack a supersonic target head-on and if you miss go round the back and have another go. It will sustain about 5 g all the way round and climb at the same time. A very exhilarating manoeuvre, but one which taxes the brain to the absolute ultimate.

BRUMAGHIM

It appears to me that the interceptor pilot is a fairly select individual. I wonder if you would give us some words about the selection process of a person able to handle such a workload. What do you feel is the most productive screening for a successful interceptor pilot?

HUTCHINSON

If you have about three or four hours I would be very happy to go into that! The interceptor pilots of the Royal Air Force are selected initially by the same means as every other pilot. They all go through the aircrew selection centre, but their training is a great deal longer in the advanced stage than that of a pilot who is destined to become a co-pilot on a bomber. I don't think anyone has discovered the ultimate in selection procedures. Inevitably selections have a wide scatter around what the desired figure would be and so selection must continue right through operational service. I would say that by delaying selection for each stage until each pilot has approached the end of the previous stage, you can reduce the error to acceptable proportions. We are always seeking to reduce the wastage, of course, because the cost of the final product is astronomical. In short we don't have an answer to the selection procedure. You cannot take a man off the street and say he will make an air defence pilot. There is no test you can subject him to because though he might well have the education, the mental and physical attributes you are looking for involve another dimension, which no one has yet managed to measure. This concerns his dynamism shall we say - the amount of energy he can apply to the job.

GUIGNARD In the kinds of mission you have been describing, particularly low level interception, can you give us a feel of how serious in practice is rough motion and vibration in interpreting displays and carrying out mental arithmetic. Is it a common or serious problem or a fairly incidental one and if it is a serious one have you any practical suggestions from a pilot's point of view as to how we might improve the situation?

HUTCHINSON I wish I could answer your question. The bumping around which goes on at low level can have a very distracting affect. To give you some feel for the problem in the case of the Lightning, which has no radio altimeter, the altimeter error at about .05 is of the order of 1,500 ft which is grounds for apprehension. There is some physical discomfort though an aircraft of that kind, which was unique when it was designed, has a very low gust response and so the ride tends to be fairly smooth. On a dark night with no horizon the time available to a pilot to interpret the radar picture is reduced. He has to devote much attention in the low level role to flying the aeroplane and so he has less of his total mental capacity to devote to interpretation. Now what can you do about it? The best thing is to put another man in the aeroplane to work the radar. We suffer, more than in any other role, when we are operating low down in the dark.

GUIGNARD Does the physical motion of being bumped around lead you to believe there is any way in which display designs may be improved?

HUTCHINSON We have a relatively low gust response in the Lightning. I think that most air defence aircraft are designed on similar lines and it does not worry me usually. If you do get into a very bumpy piece of air then you have difficulty in seeing the instruments on occasion. A head up display may help. I have only had experience of head up displays in earlier aircraft, such as the Javelin, but I found that in the darkness they tended to spoil your ability to see out.

GUIGNARD Have you or your colleagues ever experienced disruption of your mental arithmetic by very severe bumping?

HUTCHINSON The short answer to that is 'yes'.

ROLFE Could I have your comments on the value of the simulator as a device for training and maintaining skill under high workload conditions?

HUTCHINSON The simulator is a very useful tool but it has its limitations. You can load up a pilot to breaking point very easily with a simulator and to that extent it is useful for supervisors to know under what sort of load an individual is likely to crack. As a Squadron Commander I find that useful in itself. In radar interpretation a simulator is a very useful aid, but if you want to know it if helps to fly the aeroplane then the answer is 'no', because its responses are not anywhere near enough representative. In fact most of us doing simulator interceptions fly on autopilot all the time because everytime you trip the autopilot it turns upside down and dives into the sea! Similarly, there are some tasks on the existing Lightning simulator which cannot be simulated. Low level ground returns cannot be produced on the radar, so in the very high workload situation the simulator has its limits. The simulator is a very useful tool particularly to those of us who are concerned with the organisation and supervision of others, but it has very definite limitations in assessing their ability to cope with a real life situation.

TER BRAAK I fully agree with Wing Commander Hutchinson. The problem with the simulator is that it does not fly like an aircraft. You can use it for procedure training and emergency training, but you can never simulate an attack or a run in on the target. I feel that simulation of missions can only be done properly in the air in normal aircraft.

HUTCHINSON You cannot of course simulate flight refuelling in a simulator and that is one of the most testing parts of a long distance air defence mission. You cannot simulate the navigational problem of finding a tanker or of finding your combat station.

MIDDLETON I believe the maritime simulator is a good tool to determine the level at which the crew can no longer cope. I would agree with the other members that you cannot simulate all the mission profile and certainly in a complex crew structure like the maritime aeroplane you can very quickly saturate the crew members with workload.

PHYSIOLOGICAL COSTS OF EXTENDED AIRBORNE COMMAND AND CONTROL OPERATIONS

by

Ralph R. Bollinger, M.D.
Robert D. O'Donnell, Ph.D.
Bryce O. Hartman, Ph.D.
USAF School of Aerospace Medicine and
Aerospace Medical Research Laboratory
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas, USA, 78235

SUMMARY

During Exercise Night Star the personnel of the National Emergency Airborne Command Post successfully documented their ability to maintain a continuous airborne alert for an extended period. Biomedical evaluation began with a pre-exercise baseline study and continued through a postexercise observation period. A variety of psychological and physiological parameters were measured in order to determine the degree of stress, fatigue, and change in performance induced by the extended airborne alert. This biomedical evaluation showed that performance was maintained by the mission teams, flight crews, and ground support personnel. When significant fatigue did occur, whether in flight or on the ground, it developed near the beginning of the exercise. The only cases of marked or persistent fatigue were seen in those groups whose day/night, work/rest cycles were shifted and can be attributed in major part to the resulting sleep loss. However, all groups appeared to adapt to their new work schedules as the exercise progressed. Partial physiologic and complete psychologic recovery were evident within the first 36 hours after the exercise.

INTRODUCTION

Exercise Night Star was conducted in May 1973 by personnel of the National Emergency Airborne Command Post (NEACP), the Emergency Airborne element of the Joint Staff, and the first Airborne Command Control Squadron (ACCS), an elite separate squadron of the United States Air Force. As an alternate command center of the National Military Command System, NEACP provides the National Command Authorities with the emergency means essential for accurate and timely decisions for the direction of U. S. military forces. The 1st ACCS provides aircrews to man the EC-135J Airborne Command Post aircraft, as well as the necessary maintenance and communications electronics support. During the 11 years since it became operational, the capabilities of NEACP have been repeatedly expanded. Consequently, Exercise Night Star was designed to test the ability of NEACP to maintain a continuous airborne alert for an extended period. In addition to tests of equipment and procedures, an evaluation of stress, fatigue and changes in performance was also desired. For this purpose personnel from the U. S. Air Force School of Aerospace Medicine and the Aerospace Medical Research Laboratory, both constituents of the Aerospace Medical Division, AFSC, were invited to assist in the exercise. A variety of psychological and physiological parameters were measured to determine what changes were induced by the extended airborne alert. The results of that biomedical evaluation constitute this report.

STUDY PLAN

Data collection was started on 14 May 1973. Baseline information was collected from the members of one of the three mission teams over a 24-hour period while they were on duty in the alert facility. During the baseline collection period, the duty team knew they were vulnerable to the start of Exercise Night Star but were otherwise following their normal duty routines. Data were collected every 4 hours with the exception of the 0300 mid-sleep collection, which was omitted.

At 1900 hours on 15 May 1973, Exercise Night Star was initiated without warning. The alert aircraft with team 3 on board departed the alert facility in less than 5 minutes from the first sounding of the Klaxon. During the ensuing airborne alert, the mission teams were investigated intensively while less information was obtained from the flight crews and ground-support elements. Each mission team was comprised of 17 individuals occupying positions in one of five functional areas: Control Section, Intelligence, Emergency Actions and Communications, Operations and Resources, and Plans. Members of all three mission teams were subjected to the physiological and most psychological measurements shortly after takeoff and every 4 hours thereafter during all 11 flights. Most of the missions were 8½ hours long. However, the first mission of team 3 and the second mission of team 2 were 12 hours in length and included aerial refuelings. An inflight refueling was conducted during the fourth flight of team 1 as well. The group that resumed alert after Night Star was studied for 36 hours to determine recovery rates.

The results and conclusions which follow are based primarily on the findings from teams #1 and #3. The data from team #2 were consistent with that of the other teams but were not used for illustrative purposes because team #2 flew three missions while the other two teams flew four, because there was one missing data collection during each of the three missions of team #2 and because the flying times of the other two teams permitted sharper differentiation of day versus night effects.

Several comments are warranted on the physiological measurements made on the mission teams during Exercise Night Star. Flying operations have repeatedly been shown to exert considerable influence on bodily functions which are not under voluntary control. Consequently, biochemical as well as biophysical processes can be measured for the purpose of assessing flight effects, the so-called physiological "costs." These flight-induced costs are superimposed upon the rhythmic daily physiological changes that occur in all healthy persons. The baseline determinations made before the start of Exercise Night Star were therefore used in the assessment of all data collected during and following the various flights. By subtracting the baseline value in each case from the so-called "flight" or "recovery" value, it was possible to show whether the costs were normal or excessive.

The physiologic assessment was broadly based. The indicas chosen were body temperature, sodium output, potassium output, urea output, urine volume, and the balance between sodium and potassium excretion Increases in the output of these chemicals indicate that compensatory physiologic adjustments are occurring to changes induced by the exercise. Specimens were collected from each team member shortly after the start of each flight and every four hours thereafter. Thus a total of three or four specimens were collected per mission, depending upon the mission length. The early data served to show whether or not there were pre-flight elevations (predispositions), whereas the later data showed the direction and magnitude of each flight-induced physiologic shift. For each collection, the subject voided into a 250 cc plastic bottle to which 2.0 gms of boric acid had been added as a preservative. Samples were sent to the USAF School of Aerospace Medicine where laboratory analyses were performed using the AutoAnalyzer and flame photometer. Oral temperatures were taken with Yellow Springs Instruments Telethermometer probes or individually assigned mercury thermometers and were recorded by observers onboard each aircraft.

Subjective fatigue forms were filled out every four hours at approximately the same times that the physiologic sampling took place. Sleep histories were completed by each crewmember once per 24 hours following the longest sleep period. Both the fatigue and sleep history forms were those which had been used many times at the USAF School of Aerospace Medicine. Consequently, the scores obtained during Exercise Night Star could be interpreted in light of an extensive experience with a variety of stressful flying operations.

The effect of extended missions on team performance was assessed by both subjective evaluation and objective measurements. Overall changes in general characteristics such as mood, alertness, and coordination were evaluated by means of subjective rating scales, critical incident surveys, and debriefings. In addition, an attempt was made to quantify performance during certain mission segments by tape recording them and subsequently applying time and content analyses. Eight key personnel from each team were selected to fill out a rating scale twice during each mission. The team members sampled were the team chief, CCOC, operations officer, plans officer, emergency actions officer, intelligence officer, communications officer, and communications NCO (teletype). The performance evaluation was conducted at the mid-point and near the end of each flight to allow comparisons between the first and second halves of the missions. These evaluations consisted of self-rating questions, questions about specific job areas and overall mission appraisals. These scales probed the areas of response speed and accuracy, stress level, support quality, and more subtle behavioral changes such as anticipation of data requirements and decisiveness. In addition, an observer flew with each mission of team #1 and made systematic checklist observations of team activity during the mission scenarios. As a modulator of these evaluations, the team chief was debriefed after each scenario to determine the relative difficulty of that mission. Tape recordings of all intercom traffic during two of the mission scenarios of team #1 were analyzed to detect stress- or fatigue- induced changes in the team's style of operation.

In addition to the information obtained from members of the mission teams, fatigue forms and oral temperatures were obtained from the flight crews at four-hour intervals and sleep histories were filled out daily. Urine was collected every four hours from each aircraft commander. Members of the maintenance crews completed fatigue forms and sleep histories during each duty period.

PHYSIOLOGICAL RESULTS

The data on physiological costs are summarized in Table 1. Generally mild to moderate stress developed during Exercise Night Star as judged by simple metabolic indices. For example, urinary electrolyte excretion, which increases with nonspecific stress, did not rise appreciably during the flights. Correspondingly small changes are seen when pooled data for urea excretion, sodium/potassium ratio and urine volume from the flights of team 1 (day) and team 3 (night) are expressed as deviations from the baseline values. No performance decrement has been associated in past studies with this degree of physiological change. Thus, these results are compatible with the sustained crew performance actually observed.

Table 1

Physiological Cost Summary

Test Condition	Time of Day	Urinary Index* Volume	K	Na	Na/K	Urea
Preliminary	0700	43	1.9	5.5	3.5	885
	1100	67	4.0	10.3	2.6	984
	1500	84	4.0	12.2	3.3	1056
	1900	87	3.4	13.4	4.3	1183
	2300	74	2.6	14.0	6.1	1139
Flight 1	1900	85	3.0	14.4	5.1	1300
	2300	82	3.2	14.7	5.3	1322
	0300	79	2.4	10.6	4.6	1317
	0700	93	3.1	9.8	3.8	1198
Flight 2	2300	53	3.3	7.9	2.8	975
	0300	63	2.8	7.8	3.0	1029
	0700	59	2.2	7.2	3.6	1170
Flight 3	0300	54	2.7	6.9	2.9	1278
	0700	60	3.5	7.3	2.3	1564
	1100	78	4.1	9.2	2.4	1398
Flight 4	0300	43	3.1	5.5	2.2	877
	0700	76	3.5	7.1	2.4	1014
	1100	51	4.2	7.2	2.1	1040
Flight 1	0700	70	3.8	11.8	3.4	763
	1100	135	6.4	16.8	2.9	1336
	1500	62	2.9	9.5	3.9	1064
Flight 2	0700	62	4.0	8.7	2.6	1072
	1100	143	6.5	14.5	2.5	1285
	1500	75	3.8	10.6	3.0	1245
Flight 3	1100	71	6.4	10.9	1.9	1378
	1500	86	5.1	10.1	2.2	1317
	1900	106	3.7	12.6	4.0	1683
Flight 4	1100	73	5.3	7.8	1.5	1055
	1500	66	3.7	7.6	2.1	932
	1900	74	3.4	10.0	3.4	1329
Recovery	0700	57	2.4	5.3	2.2	1192
	1100	57	4.9	7.1	1.5	1130
	1500	109	5.1	9.5	1.9	1164
	1900	81	3.6	9.7	2.9	1241
	2300	73	3.1	9.9	3.2	1320
	0700	64	2.0	7.1	3.8	1436

*Each index is a creatinine-based ratio (ml, mEq, or mg/100 mg creat.)

Although the effects are mild, they demonstrate a difference in the responses of the team depending on whether they flew at night or during the day. For example Table 2 shows sodium output as milliequivalents of deviation from baseline values. The day team showed a relatively high sensitivity (indicated by larger positive values) to flight factors early in the mission whereas the night team consistently showed it at the late time. This table also illustrates another interesting finding. Flights 1 and 2 were the most stressful for both teams as indicated by sodium excretion. This finding is confirmed by the other urinary measurements. Either the workload decreased or adaptation occurred, resulting in decreased physiologic cost for later flights. A corollary of this finding is that cumulative stress evidently did not become a factor over the 96-hour exercise. The minimal changes observed during the operation indicate that the work schedule could have been maintained for additional missions if necessary.

Table 2

PHYSIOLOGICAL COST

SODIUM OUTPUT*

Flight	Early in Flight Day	Early in Flight Night	Late in Flight Day	Late in Flight Night
#1	+6.3	+0.7	-2.7	+4.3
#2	+3.2	-6.1	-1.6	+1.7
#3	+0.6	-2.9	-0.8	-1.1
#4	-2.5	-4.3	-3.4	-3.1

*In mEq of deviation from baseline values

During the 36-hour postexercise observation period, only partial physiological recovery was detected. For example, in Figure 1, urea excretion in mg per 100 mg of creatinine is plotted against time of day for the baseline period, the flying portions of the exercise and the recovery period. Although the recovery values are reapproaching the baseline values, some flight induced deviations are still evident.

Oral temperature data reveal that persistent hyperthermia was present during all flights of Exercise Night Star. The temperature elevation was generally greater in the case of team 3 flying at night than that of team 1 flying during the day. In Figure 2, the pooled data from all three teams substantiates the progressive fall in body temperature between the beginning and end of each mission previously observed by Harris et al. (1). This overall temperature change paralleled the decrease in alertness (increased fatigue) detected using the subjective fatigue forms.

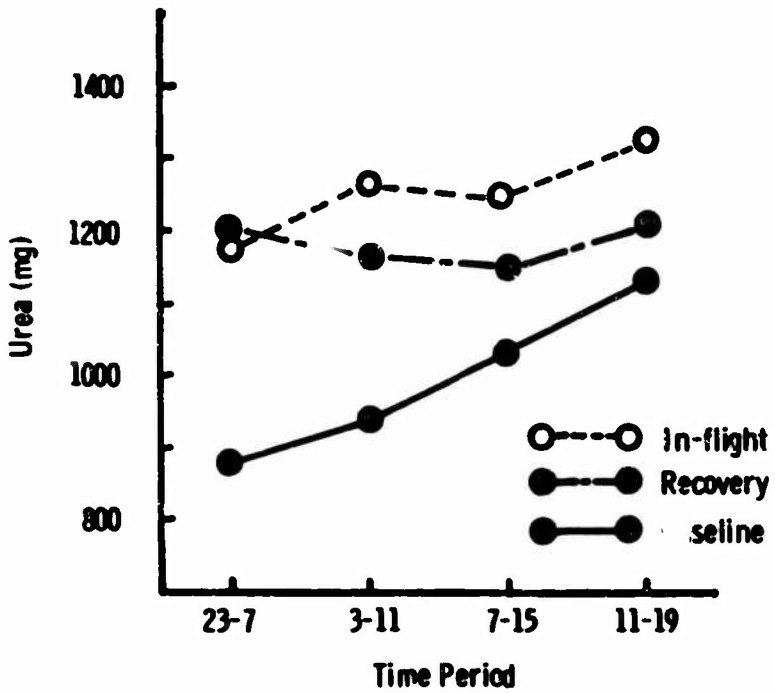


Figure 1: Physiological Cost - Urea Excretion

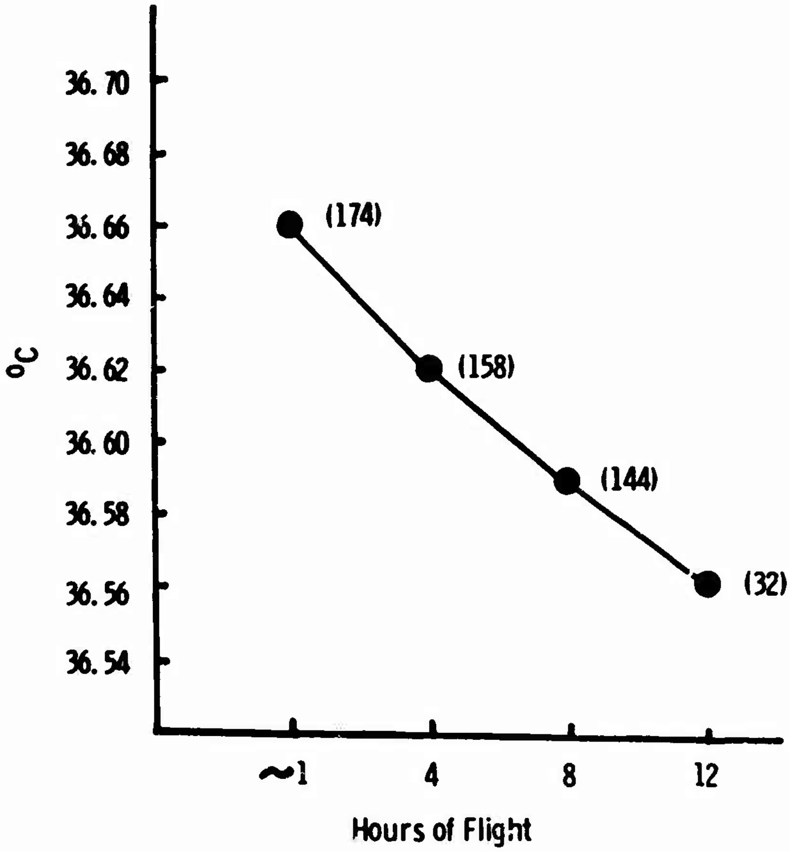


Figure 2: Oral Temperature

PSYCHOLOGICAL RESULTS

Exercise Night Star generally produced only mild fatigue, although flying always produced significantly greater fatigue than that occurring normally during the same periods on the control day. However, with the exception of the first 12-hour mission performed by team 3, the fatigue level seldom fell below a mean score of 9 (lower scores indicate increasing fatigue). Previous studies have found that complete recovery following an 8 hour sleep period normally occurs when fatigue is at this level.

The maximum mission team fatigue occurred at the end of the first 12-hour flight (Figure 3). This level of fatigue (mean subjective fatigue score at 0700 hours, 16 May for team 3 = 6.7) has previously been found to produce incomplete recovery within one normal sleep period. This relatively low score was probably a result of the combined effects of the 12-hour mission, plus the preceding 12 hours of alert duty. The magnitude of the fatigue in this case was at a level consistent with moderate persistent psychological cost, although performance was maintained at an acceptable level.

Figure 3 also shows that the team 3 average early mission fatigue increased over missions. This may be the result of incomplete recovery from mission 1. However, it should also be noted that mean late mission scores were similar for all missions indicating that overall fatigue was mild.

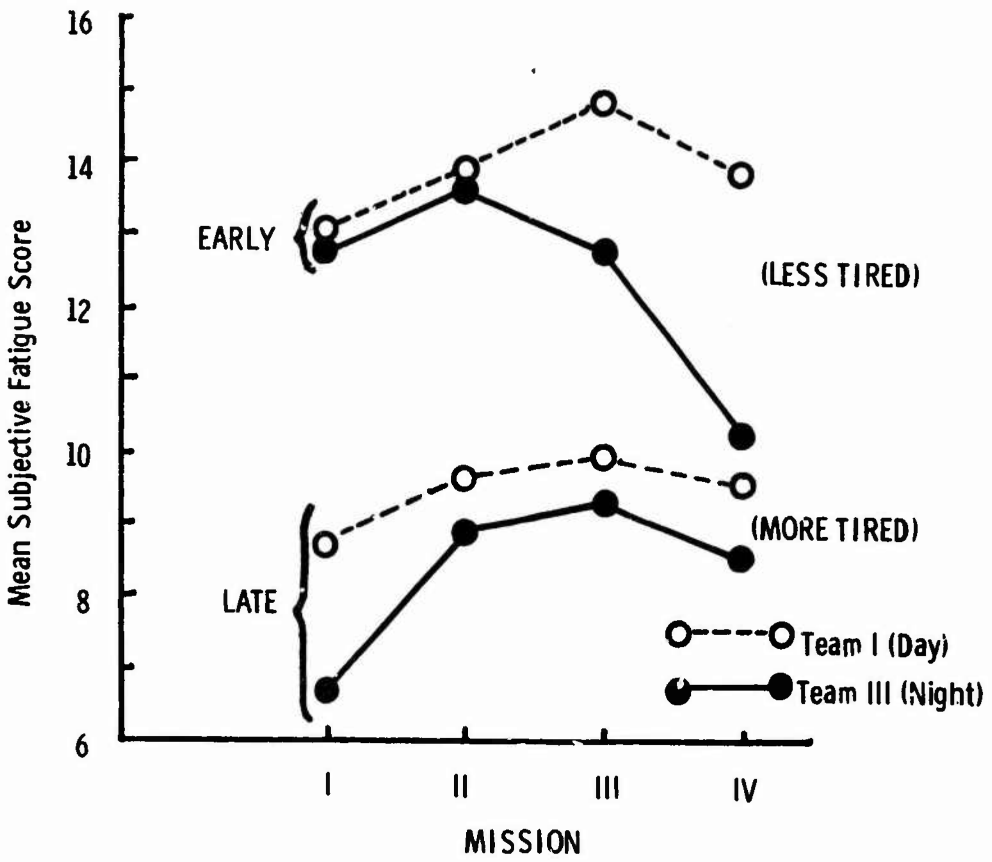


Figure 3: Subjective Fatigue Early vs Late in Flights

There was little evidence of cumulative fatigue over missions. A comparison of early mission, mid-mission and late mission scores on mission 4 and mission 1 (team 1) and late mission scores on mission 4 and mission 2 (team 3), all corrected for time of day, showed only one significant difference. Team 1 showed relatively greater early mission fatigue on mission 1 than mission 4. The only evidence for any cumulative fatigue was the progressive decrement in early mission scores for team 3. However, this may be due in part to the small number of hours (5.6) devoted to sleeping prior to their fourth mission.

The mean number of hours of sleep obtained during Exercise Night Star are shown in Table 3. There was no sleep loss for team 1 which performed four daytime missions and slept at night. On the other hand, the members of team 3, who were required to sleep during the day, slept an average of 1.2 hours less than team 1. Consequently, there was a difference in the responses of team 1 and team 3 to the question "How do you feel?" Team 1 personnel awoke feeling refreshed each day, whereas the members of team 3 felt somewhat less rested. Sleep loss was responsible, at least in part, for the progressive decrement in the early mission scores of team 3. One way to detect cumulative effects is to compare recovery data with pre-exercise base-line data. Figure 4 shows that recovery from fatigue appeared to be complete within 24 hours.

Table 3

MEAN NUMBER HOURS OF SLEEP

OBTAINED DURING EXERCISE NIGHT-STAR

PHASE	TEAM 1	TEAM 3
Pre-exercise Control	6.2	6.5
Between Missions 1 and 2	6.8	6.0
Between Missions 2 and 3	8.2	7.5
Between Missions 3 and 4	8.0	5.6
Postexercise Recovery	7.3	

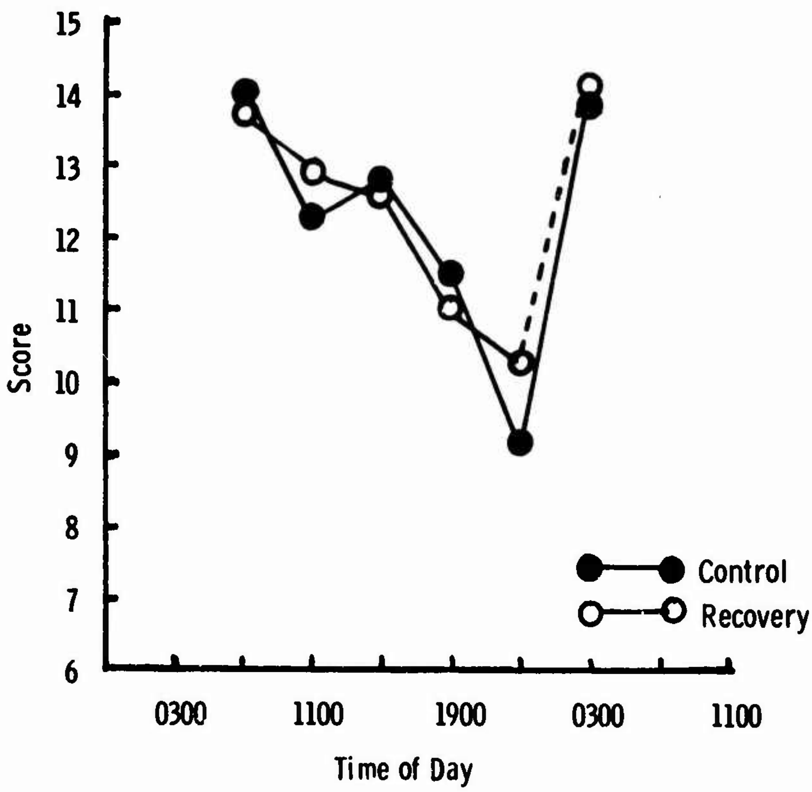


Figure 4: Subjective Recovery om Fatigue

Performance ratings were obtained by administering a questionnaire to 8 key individuals on the mission teams. Although about 16% of the data is missing, the responses are internally consistent and seem to represent valid measures of the feelings and attitudes of the respondents. No one failed to make at least some differentiation in the various estimates. Observational data also appear reliable, but suffer from the lack of standardization between missions.

Subjective performance efficiency was determined from the subject's own impression of his performance as indicated by the number of mistakes he noted in details of the mission and by people working for him. There was no significant difference between halves of the missions and there was little change over the four missions in subjective performance efficiency. These measures were interpreted by the subjects as indications of how well things were actually going during the flights. They showed that at least as far as the team members were concerned there was no decrement in performance during the exercise.

However, respondents seemed less certain of their work as the test period progressed. They reported working somewhat slower and feeling slightly less "accurate" with succeeding missions. The overall impression is that real efficiency and capability to react were not impaired over the test period. However, by the end of the period, more than normal care and attention were required to maintain efficiency. Considered in light of the overall feeling that just as good a job was being done, the decreased certainty probably represents a slight change in the style of working. Such a minor performance change would be of no operational significance.

Flight Crew Subjective Fatigue

Thus far, the findings from the mission teams have been discussed since they provided the bulk of the biomedical data. The information on flight crews is sparse and difficult to analyze since few personnel were involved and fewer missions were flown per person. Furthermore, the missions flown were often on different shifts, thus making individual comparisons impossible. However, a few comments are warranted. As shown on Figure 5, the overall mean subjective fatigue score after 8 hours of flying was 9.6, an acceptable level. However, the aircraft commander, co-pilot and navigator from the first mission of the exercise represent a significant exception. They were required to fly a 12-hour night mission including an inflight refueling following a full work day on duty in the alert facility. Their mean score was 3.0, a level of fatigue which can be associated with significant performance decrement. The flight crew evaluations also showed that there was little variation among crew positions in subjective fatigue after 8 hours in flight. However, a comparison of average scores for crews flying 8-hour overnight missions with those of crews flying during the day, showed that night flights produced greater fatigue (mean score 7.6) than daytime flights (mean score 10.5).

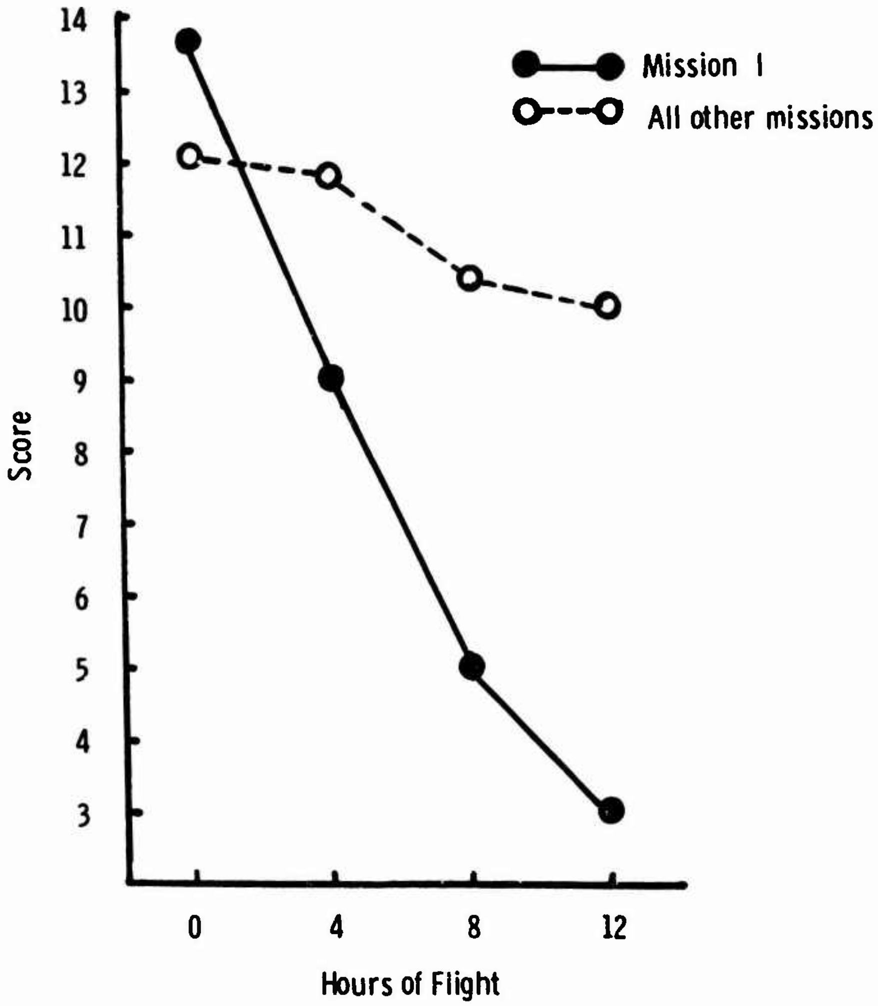


Figure 5: Mean Subjective Fatigue in Flightcrews

Ground Crew Subjective Fatigue

Although the data on the ground support elements is also limited, several conclusions seem justified. The fatigue scores of night crews at the end of the first two work shifts were 4.0 and 5.4, a level of fatigue which can be associated with decreased efficiency. The low values in the first two nights are probably due primarily to the abrupt change from day to night work which most night shift workers experienced when the alert began. There appeared to be a slight recovery from fatigue by the end of the exercise. An apparent progressive increase in feelings of fatigue was detected among daytime personnel. However, other data shows that this change occurred primarily in crews working a 12-hour shift. Those on an 8-hour day shift showed no change.

CONCLUSIONS

The biomedical evaluation conducted during Exercise Night Star has revealed the following findings:

1. The general level of stress and fatigue was mild.

2. Performance was maintained by the mission teams, flight crews, and ground support personnel

3. When significant fatigue did occur whether in flight or on the ground, it occurred during the first part of the exercise rather than near the end.

4. The greatest effects were seen in persons who had to work at night both in the air and on the ground. The only cases of marked or persistent fatigue occurred in those groups whose day-night work/rest cycles were shifted. This can be attributed in major part to the resulting sleep loss.

5. All groups appeared to adapt to their new work/rest schedules as the exercise progressed.

6. No cumulative effects were seen.

7. Partial physiologic and complete psychologic recovery were evident within the first 36 hours after the exercise.

Operational Applications

The biomedical findings from Exercise Night Star have several immediate and future operational applications. No crew vulnerabilities were revealed which would require immediate changes in the mode of NEACP operations. Airborne alerts for periods longer than Exercise Night Star appear biomedically feasible. Finally, a data base has been accumulated and personnel subsystem test procedures have been developed for the initial operational test and evaluation of the E-4A Advanced Airborne Command Post.

ACKNOWLEDGEMENT:

The invaluable assistance of Sgt. Cordell Gardiner during the field phase of this study is gratefully acknowledged. Mr. Robert Miranda and Mr. Edgar Williams performed the chemical analyses. Drs. James Ellis and James Sanford contributed significantly to our evaluation of the data. Mr. Richard McNee performed the statistical analyses.

REFERENCE:

1. Harris, D. A., H. B. Hale, B. O. Hartman, and J. A. Martinez. Oral temperature in relation to inflight work/rest schedules. Aerospace Medicine 41:723, 1970.